Engineering Applications of Artificial Intelligence 126 (2023) 107098

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A novel shape-based averaging algorithm for time series

Yutao Liu [a,*], Yong-An Zhang [b], Ming Zeng [b], Jie Zhao [a]

[a] The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China
[b] Harbin Institute of Technology, Harbin 150001, Heilongjiang, China

ARTICLE INFO

Keywords:
Time series averaging
Similarity measure
Dynamic time warping
Time series classification
Template generation

ABSTRACT

Time series averaging is one of the essential subroutines in time series analysis. DTW Barycenter Averaging (DBA) has proven to be an effective and popular DTW-based time series averaging algorithm. However, DBA lacks the ability to average time series in the time domain, making it sensitive to initialization. In this research, we propose a novel shape-based time series averaging algorithm, called Shape DTW Weighted Averaging (ShapeDWA), to address the shortcomings of DBA. The proposed ShapeDWA algorithm combines the advantages of the DBA and the Cubic-spline DTW (CDTW) averaging methods. The concepts of time index averaging and re-sampling in the CDTW algorithm are incorporated into the DBA algorithm, giving ShapeDWA the ability to average a set of time series in both the amplitude and time domains. Moreover, ShapeDWA utilizes a weighed average instead of the barycenter average in DBA, which effectively attenuate the effects of noise, outliers, and local amplitude differences between the time series. To qualitatively evaluate and compare the proposed time series averaging algorithm, two metrics have been developed: average discrepancy distance and average time distortion. Extensive experimental results on the UCR time series database illustrate the superior performance of ShapeDWA over DBA and SSG, with an average reduction of 23.42% and 24.89% for average discrepancy distance, and 18.76% and 19.81% for average time distortion. Furthermore, the template matching-based classification experiment shows that ShapeDWA combined with these two developed metrics improves the classification rate by 17.07% and 16.42% compared to DBA and SSG, respectively.

1. Introduction

Time series are sequences of time-dependent observations that are currently ubiquitous and have been applied in fields as diverse as climate (Mudelsee, 2019), clinical medicine (Salgotra et al., 2020), agriculture (Maus et al., 2016; Belgiu and Csillik, 2018), biological science (Redmond et al., 2020), financial engineering (D'Urso et al., 2021), etc. Time series averaging plays an indispensable role in some data analysis tasks, including classification (Petitjean et al., 2016; Tran et al., 2019; Okawa, 2021), clustering (Petitjean et al., 2011; Izakian et al., 2015; Łuczak, 2016), noise suppression (Kotas et al., 2015; Zhu and Najafizadeh, 2017) and information retrieval (Dau et al., 2019). For example, using the average time series as the templates can significantly improve the performance of the template matching-based classifier without increasing the computational complexity (Liu et al., 2018; Okawa, 2019, 2020). Some clustering methods, such as K-Means clustering (Petitjean et al., 2011; Sun et al., 2016), fuzzy clustering (Izakian et al., 2015) and hierarchical clustering (Łuczak, 2016), also require an averaging method.

Time distortions between time series, such as shift, scaling, missing and redundancy, make averaging time series a challenging problem (Petitjean et al., 2011). Dynamic time warping (DTW) is a well-known method for sequence alignment and similarity measurement between time series because it is able to capture flexible similarities under time distortions (Sakoe and Chiba, 1978). Time series averaging under DTW provides a potential solution. However, since the DTW distance is not a standard metric, a polynomial-time algorithm for computing an exact average time series is unknown, except for some special time series, such as binary time series (Petitjean et al., 2011; Brill et al., 2019). Therefore, a number of heuristic methods have been developed to obtain an approximate solution with acceptable computational complexity.

DTW Barycenter Averaging (DBA), developed by Petitjean et al. has been demonstrated to be the most effective time series averaging algorithm (Petitjean et al., 2011, 2014, 2016). DBA is a global optimization based averaging algorithm that iteratively refines the average time series to minimize the average DTW distance. Unfortunately, DBA has two fatal shortcomings, as shown in Figs. 2 and 8. First, it is

* Corresponding author.
E-mail addresses: liuyutao190160@126.com (Y. Liu), zhangyongan76@163.com (Y.-A. Zhang), zengming@hit.edu.cn (M. Zeng), zhaojie@zzu.edu.cn (J. Zhao).

https://doi.org/10.1016/j.engappai.2023.107098
Received 8 October 2022; Received in revised form 26 July 2023; Accepted 1 September 2023
Available online 15 September 2023

sensitive to the initialization, especially for the time series dataset of different lengths or with large time distortions (Liu et al., 2019; Sioros and Nymoen, 2021). As a result, the length and the spatio-temporal characteristics of the resulting average time series are highly dependent on the initial average time series. Secondly, there are some spatio-temporal distortions in the resulting average time series, especially near extreme points. These two shortcomings are mainly due to the fact that DBA lacks the ability to average the time series in the time domain.

A desired average time series should contain the dominant spatio-temporal characteristics of the time series dataset (Liu et al., 2019). For instance, a motion template of moderate amplitude and velocity is required as a reference in the assessment of human motor function, rehabilitation training, gymnastics, dance, etc (Hachaj et al., 2017). In other words, the time series averaging algorithm should be able to average a set of time series in both the amplitude and time domains. In Niennattrakul et al. (2012), a Cubic-spline dynamic time warping (CDTW) averaging function was developed to average two time series in both the amplitude and time domains. Inspired by the CDTW and DBA methods, we proposed a novel shape-based averaging algorithm, called Shape DTW Weighted Averaging (ShapeDWA). The main contributions of this paper are summarized as follows:

- (1) The strategies of time index averaging and re-sampling in the CDTW algorithm are introduced into the DBA algorithm. As a result, ShapeDWA is able to average a group of time series in both the amplitude and time domains, resulting in an average time series that preserves the spatio-temporal characteristics. Experimental results also show that ShapeDWA can effectively solve the problem that DBA is sensitive to initialization.
- (2) The barycenter average in DBA is replaced by a weighted average, where the weight of each element in the time series depends on the alignment between it and the temporary average time series. This effectively reduces the effects of noise, outliers, and local amplitude differences in the time series.
- (3) Two metrics, average discrepancy distance and average time distortion, are introduced to evaluate the performance of the proposed ShapeDWA algorithm compared to other algorithms.

The remainder of this paper is organized as follows. Section 2 reviews some preliminaries, including the DTW algorithm, time distortion measure, time series averaging under DTW and the DBA algorithm. In the following section, the proposed ShapeDWA algorithm is presented. In Section 4, we first point out the inadequacies of the DTW distance and then introduce two metrics to assess the quality of the average time series. The experimental results are presented and discussed in Section 5. Finally, the conclusions and further work are presented in Section 6.

2. Background and related work

In this section, we provide some essential background knowledge and related work to understand our proposed method.

2.1. Dynamic time warping

A time series $\boldsymbol{X}$ is an ordered sequence of real numbers: $\boldsymbol{X} = \left[x_1, x_2, \dots, x_n\right]$. x_i can be univariate or multivariate. This paper focuses on the univariate time series. Usually, the sampling period is constant, so the time information is replaced by its index. In addition to amplitude differences, there are time distortions between time series. DTW is an effective method for sequence alignment and similarity measurement between time series. It uses dynamic programming to find the optimal alignment with minimum cumulative distance.

Given two time series $\boldsymbol{P}$ and $\boldsymbol{Q}$, of length m and n respectively, where:

$\boldsymbol{P} = \{p_1, p_2, \dots, p_m\}$

$\boldsymbol{Q} = \{q_1, q_2, \dots, q_n\}$

The alignment between two time series can be represented by an $L \times 2$ matrix $\boldsymbol{W}$, also called warping path. Each row vector $\boldsymbol{w}_k = (i,j) \in [1:m] \times [1:n]$ for $k \in [1:L]$ of the matrix $\boldsymbol{W}$ represents a pair of alignment, where i and j denote the index of elements in $\boldsymbol{P}$ and $\boldsymbol{Q}$, respectively.

Considering the time-order of time series, the alignment is typically subjected to the following constraints:

- Boundary constraint: the first and last points of time series must be aligned with each other, i.e., $\boldsymbol{w}_1 = (1,1)$ and $\boldsymbol{w}_L = (m,n)$.
- Step size constraint: the transition from one point to the next point must along the following directions: diagonal, up, and right, i.e., $\boldsymbol{w}_{k+1} - \boldsymbol{w}_k \in \{(1,1),(1,0),(0,1)\}$ for $k \in [1:L-1]$. It means that the warping path cannot go back in time, cross, or make any jumps.

According to these above constraints, the length L of a warping path satisfies the following condition:

$\max(m,n) \le L < m+n-1$

For the warping path $\boldsymbol{W}$, the cumulative distance between $\boldsymbol{P}$ and $\boldsymbol{Q}$ is defined as follows:

$$d_{\boldsymbol{W}}(\boldsymbol{P},\boldsymbol{Q}) = \sum_{k=1}^{L} d\left(p_{w_{k,1}}, q_{w_{k,2}}\right) \tag{1}$$

where $p_{w_{k,1}}$ and $q_{w_{k,2}}$ are kth pair of aligned elements in $\boldsymbol{P}$ and $\boldsymbol{Q}$ respectively. $d(\cdot)$ denotes the local distance function between two aligned elements. For the classical DTW algorithm, the squared Euclidean distance is utilized as the local distance.

The DTW warping path $\boldsymbol{W}^*$ is defined as a warping path having the smallest cumulative distance, and the corresponding cumulative distance is referred to as the DTW distance,

$$\begin{aligned} d_{DTW}(\boldsymbol{P},\boldsymbol{Q}) &= d_{\boldsymbol{W}^*}(\boldsymbol{P},\boldsymbol{Q}) \\ &= \min\{d_{\boldsymbol{W}}(\boldsymbol{P},\boldsymbol{Q}) | \boldsymbol{W} \in \mathbb{W}_{m,n}\} \end{aligned} \tag{2}$$

where $\mathbb{W}_{m,n}$ is the space of all possible warping paths of two time series with length m and n. It should be noted that DTW warping path is not unique, but the DTW distance is unique.

The optimal warping path can be efficiently discovered using the dynamic programming technique. An $m \times n$ cumulative distance matrix $\boldsymbol{D}$ is constructed, where the element $\boldsymbol{D}_{i,j} \in \mathbb{R}^{i \times j}, 1 \le i \le m, 1 \le j \le n$ is defined as follows:

$$D_{i,j} = \begin{cases} d_{i,j} & i=1, j=1 \\ d_{i,j} + D_{i,j-1} & i=1, j>1 \\ d_{i,j} + D_{i-1,j} & i>1, j=1 \\ d_{i,j} + \min\begin{cases} D_{i-1,j-1} \\ D_{i-1,j} \\ D_{i,j-1} \end{cases} & i>1, j>1 \end{cases} \tag{3}$$

where $d_{i,j}$ is the local distance between p_i and q_j.

Based on Eq. (3), we fill in the matrix $\boldsymbol{D}$ from left to right and from bottom to top. The upper-right cell $D_{m,n}$ gives the total cumulative distance, i.e., $d_{DTW}(\boldsymbol{P},\boldsymbol{Q}) = D_{m,n}$. Using the backtracking procedure, the DTW warping path can be revealed from the cumulative matrix based on the opposite direction of the step constraint. The pseudo code of the calculation of DTW warping path is presented in Algorithm 1.

Even if the local distance function is a metric, the DTW distance is generally only a pseudo-simi metric satisfying:

- $d_{DTW}(\boldsymbol{P},\boldsymbol{Q}) = d_{DTW}(\boldsymbol{Q},\boldsymbol{P})$ (symmetric).
- $d_{DTW}(\boldsymbol{P},\boldsymbol{Q}) \ge 0$ and $d_{DTW}(\boldsymbol{P},\boldsymbol{P}) = 0$, but $d_{DTW}(\boldsymbol{P},\boldsymbol{Q}) = 0 \not\Rightarrow \boldsymbol{P} = \boldsymbol{Q}$.

Algorithm 1 DTW warping path

Input: Cumulative distance matrix $\boldsymbol{D}$.
Output: DTW warping path: $\boldsymbol{W}^*$.

Initialization: $i = m, j = n, \boldsymbol{W}^* = [\ ]$;
while $(i > 1) \| (j > 1)$ **do**
  if $j == 1$ **then**
    $i = i - 1$;
  else
    $i = i - \left(D_{i-1,j} \leq D_{i,j-1} \| D_{i-1,j-1} \leq D_{i,j-1}\right)$;
  end if
  if $i == 1$ **then**
    $j = j - 1$;
  else
    $j = j - \left(D_{i,j-1} \leq D_{i-1,j} \| D_{i-1,j-1} \leq D_{i-1,j}\right)$;
  end if
  $\boldsymbol{W}^* = \left[\boldsymbol{W}^*; i, j\right]$;
end while
return $\boldsymbol{W}^*$.

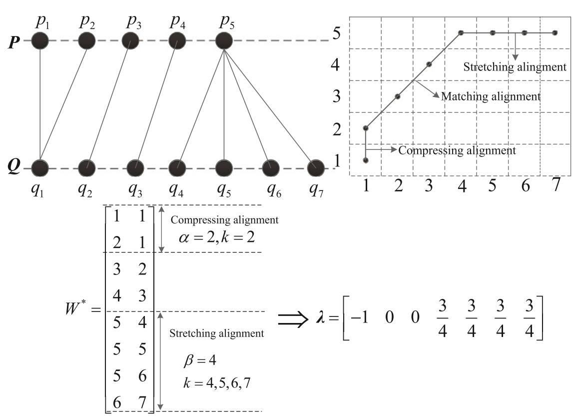

Fig. 1. Three types of DTW alignment and the calculation of time distortion coefficient.

2.2. Time distortion measure

Although the DTW distance is widely used as a measure of similarity between time series, it lacks the ability to evaluate the similarity from a temporal perspective. This even leads to errors when evaluating the similarity between time series (Folgado et al., 2018). As an elastic measure, DTW supports elastic alignments such as matching alignment, compressing alignment and stretching alignment, as shown in Fig. 1. Therefore, the time distortions between time series can be extracted from DTW warping path.

The matching alignment is a one-to-one alignment, and the warping path follows the diagonal direction, i.e., $\boldsymbol{w}_{k+1} - \boldsymbol{w}_k = (1,1)$. The compressing alignment is a one-to-many alignment, where a single point q_i in $\boldsymbol{Q}$ aligns with α consecutive points $\{p_j, p_{j+1}, \ldots, p_{j+\alpha-1}\}$ in $\boldsymbol{P}$. The warping path follows the vertical direction, i.e., $\boldsymbol{w}_{k+1} - \boldsymbol{w}_k = (1,0)$. In contrast, the stretching alignment is a many-to-one alignment, where β consecutive points $\{q_i, q_{i+1}, \ldots, q_{i+\beta-1}\}$ in $\boldsymbol{Q}$ align with single point p_j in $\boldsymbol{P}$. The warping path follows the horizontal direction, i.e., $\boldsymbol{w}_{k+1} - \boldsymbol{w}_k = (0,1)$. According to the type of alignment, the time distortion coefficient for each element of the test time series is defined as:

$$\lambda(k) = \begin{cases} 0 & \text{For matching alignment} \\ 1-\alpha & \text{For compressing alignment} \\ 1-\frac{1}{\beta} & \text{For stretching alignment} \end{cases} \tag{4}$$

The time distortion coefficient provides comprehensive time distortion information between time series, including the type and degree of time distortion at each point. The sum of the absolute value of time distortion coefficient at each point can be utilized to quantify the total time distortion between two time series,

$$\rho(\boldsymbol{P}, \boldsymbol{Q}) = \sum_{k=1}^{n} |\lambda(k)| \tag{5}$$

Similar to the DTW distance, the time distortion measure is also a pseudo-simi metric. Since the compressing alignment and the stretching alignment between two time series are reciprocal, symmetry is satisfied, i.e., $\rho(\boldsymbol{P}, \boldsymbol{Q}) = \rho(\boldsymbol{Q}, \boldsymbol{P})$. The absolute value sign in Eq. (5) determines that the time distortion measure is non-negative, i.e., $\rho(\boldsymbol{P}, \boldsymbol{Q}) \geq 0$. However, we cannot obtain $\boldsymbol{P} = \boldsymbol{Q}$ from $\rho(\boldsymbol{P}, \boldsymbol{Q}) = 0$.

2.3. Time series averaging

2.3.1. Time series averaging under DTW

Unlike the mean in Euclidean space, time series averaging under DTW is a challenging task because the DTW distance is not a standard metric. Currently, a polynomial-time algorithm for computing an exact average time series is unknown (Petitjean et al., 2011; Brill et al., 2019). Therefore, various heuristic methods have been developed to obtain a sub-optimal solution. These methods can be classified into three categories: (1) multiple sequence alignment averaging, (2) iterative pairwise averaging and (3) global optimization averaging, as summarized in Table 1.

The Steiner theory and several works in computational biology converts time series averaging under DTW into multiple sequence alignment problem (Needleman and Wunsch, 1970; Gusfield, 1997; Petitjean et al., 2011; Petitjean and Gançarski, 2012). Multiple sequence alignment averaging is a straightforward method. First, the time series are aligned all together by extending DTW. Then, the average time series can be computed by averaging column-by-column the multiple alignment. Unfortunately, the multiple sequence alignment is a typical NP-hard problem that requires unrealistic computational complexity (Wang and Jiang, 1994; Just, 2001). As a result, only a small amount of literature focuses on this approach. In Petitjean and Gançarski (2012) Petitjean and Gançarski developed a genetic based averaging method called Compact Multiple Alignment for Sequence Averaging (COMASA). In addition, even multiple sequence alignment averaging does not guarantee to provide an optimal solution (Brill et al., 2019).

Iterative pairwise averaging approaches iteratively merge two time series into an average time series in a predefined order until a single average time series remains. The main differences between these iterative averaging algorithms are the order in which the time series are pairwise averaged, and the strategy used to compute an average of two time series (Petitjean et al., 2011; Schultz and Jain, 2018). Nonlinear Alignment and Averaging Filter (NLAAF) (Gupta et al., 1996) and Prioritized Shape Averaging (PSA) (Niennattrakul and Ratanamahatana, 2009; Meesrikamolkul et al., 2012) are two representative methods. NLAAF utilized the tournament scheme to arrange the average order (Gupta et al., 1996). As the time series are randomly paired and averaged, NLAAF is sensitive to the order of averaging, with no guarantee that a different order of averaging would produce the same result. Therefore, PSA employed an ascendant hierarchical scheme, where the time series are arranged in order of similarity (Niennattrakul and Ratanamahatana, 2009). In this way, the most similar time series are aligned and averaged first. Since NLAAF and PSA utilize the association averaging strategy, each pairwise averaging operation almost doubles the length of the averaged time series, resulting in a much longer averaged time series compared to the original time series (Petitjean et al., 2011). Therefore, some re-sampling (Niennattrakul et al., 2012) and uniform scaling methods (Niennattrakul and Ratanamahatana, 2009; Fu et al., 2008) are required to reduce the length of the resulting averaged time series. In Niennattrakul et al. (2012), a shape-based averaging function called Cubic-spline dynamic time warping (CDTW) was developed to average two time series in both the amplitude and time domains. The

Table 1
Summarization of DTW-based time series averaging methods.

Method	Type	Characteristics and drawbacks
COMASA	(1)	Genetic based algorithm; large amounts of calculation
NLAAF	(2)	Tournament scheme; order sensitivity
PSA	(2)	Ascendant hierarchical scheme; order sensitivity
Shape average	(2)	Time index averaging, re-sample, ascendant hierarchical scheme; order sensitivity
DBA	(3)	Barycenter averaging, monotonically convergence; initialization sensitivity
SSG	(3)	Stochastic subgradient optimization; not descent, initialization sensitivity, parameter sensitivity
$m\_d(dp)^2$	(2),(3)	Hybrid, time series clustering; parameter sensitivity
ShapeDWA	(3)	Spatio-temporal averaging, weighted averaging, initialization robustness, parameterless

indexes of the aligned elements are also averaged as the timestamps of the average time series. A cubic-spline interpolation is then utilized to re-sample the non-uniformly sampled average time series. Based on the CDTW and the ascendant hierarchical scheme, a shape-based averaging method has been proposed to generate characteristic-preserving templates.

For global optimization averaging methods, time series averaging under DTW is formulated as an optimization problem. Given a set of N time series $\mathbb{S} = \{\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_N\}$. Without loss of generality, these time series are of different lengths. The average time series $\bar{\boldsymbol{A}}$ of $\mathbb{S}$ is defined as the time series with the minimum average DTW distance between $\bar{\boldsymbol{A}}$ and $\mathbb{S}$:

$$\bar{\boldsymbol{A}} \triangleq \arg\min \frac{1}{N} \sum_{j=1}^{N} d_{DTW}(\boldsymbol{A}, \boldsymbol{S}_j) \tag{6}$$

The global optimization averaging method has the potential to achieve a sub-optimal result with an acceptable amount of computation. For this reason, it has attracted increasing attention in recent years. Several methods, such as DTW Barycenter Averaging (DBA) (Petitjean et al., 2011) and Stochastic Subgradient Mean (SSG) method (Schultz and Jain, 2018), have been proposed. Since these two methods require an initial average time series to launch the iterative optimization process, they are sensitive to initialization, as shown in Figs. 2 and 8. The main difference lies in the optimization strategy, with DBA using barycenter average optimization and SSG using stochastic subgradient optimization. Note that DBA is a monotonically convergence algorithm which SSG is not a descent algorithm. Global Constrained Degree-pruning Dynamic Programming $m\_d(dp)^2$ (Sun et al., 2016) is a hybrid algorithm that combines global optimization averaging and iterative pairwise averaging. First, the time series are divided into M subsets. For each subset, the central time series is then calculated by DBA. Finally, M central time series are merged into the final average time series by iterative pairwise averaging. Both SSG and $m\_d(dp)^2$ require manually selected parameters. However, the selection of the parameter is also a difficult problem.

2.3.2. DTW barycenter averaging

DBA is an iterative optimization method that repeatedly refines the temporary average time series to minimize the average DTW distance. Starting with an randomly selected initial average time series, the temporary average time series is optimized iteratively by the following two steps:

- (1) Finding the alignments between the temporary average time series and each time series using the DTW algorithm.
- (2) Updating the element of the temporary average time series using the barycenter average of the elements in each time series that align with it.

Algorithm 2 details the procedure of DBA. DBA is a monotonically convergence method. The iteration process will terminate when the maximum number of iterations is exceeded or when it converges to a certain value. As mentioned above, since DBA lacks the ability to average time series in the time domain, the average time series is sensitive to initialization and has spatio-temporal distortions. To address the issue of DBA being sensitive to initialization, Okawa developed an initialization method to generate initial average time series (Okawa, 2021, 2020). In this method, all time series are re-sampled to the average length and then averaged in the Euclidean space as the initial average time series. However, this method ignores the time distortions between time series and even yields a worse result.

Algorithm 2 DBA($\mathbb{S}$, $\bar{\boldsymbol{A}}_0$)

Input: The target sequence set $\mathbb{S} = \{\boldsymbol{S}_1, \cdots, \boldsymbol{S}_N\}$, the initial average sequence $\bar{\boldsymbol{A}}_0 = \boldsymbol{S}_n, n \in [1:N]$, and the convergence threshold ϵ.
Output: Average sequence $\bar{\boldsymbol{A}}$

```
1: for i = 1 : I do
2:     Step 1: Align A_i and S_j, j = 1,2,⋯,N through DTW.
3:     Step 2: Identify the elements in S_j, j = 1,2,⋯,N that align with
       ā_i^k ∈ Ā_i, k = 1,2,⋯,l
       ā_i^k ← C_k = [c_1,⋯,c_{γ_k}]
4:     Step 3: Update each element of Ā_i by computing the average of all the
       aligned elements associated with it
       ā_{i+1}^k = (1/γ_k) Σ_{j=1}^{γ_k} c_j
5:     if d_D(Ā_i, S) − d_D(Ā_{i+1}, S) < ε then
6:         break
7:     end if
8: end for
9: return Ā = Ā_{i+1}, d'_D(Ā, S)
```

3. Shape-based time series averaging method

In this section, we propose a novel shape-based time series called ShapeDWA to overcome the shortcomings of DBA. The proposed ShapeDWA algorithm combines the advantages of DBA and CDTW, which is able to average a set of time series in both the amplitude and time domains, as shown in Figs. 2 and 8. As a result, the averaging time series calculated by ShapeDWA can preserve the spatio-temporal characteristics. Similar to DBA, ShapeDWA requires an initial average time series that is randomly selected from the dataset. For each iteration, ShapeDWA works in the following steps:

- (1) Finding the alignments between the temporary average time series and each time series using the DTW algorithm.
- (2) Updating each point of the average time series as the weighted average of the time indexes and amplitudes of the points in all time series that align with it.
- (3) Re-sampling the updated temporary average time series to the original sampling rate using the Modified Akima Cubic Hermite Interpolation (Makima) method.

Step (1) is similar to DBA. There are two innovations in step (2) compared to DBA. First, the time index averaging strategy in CDTW is introduced to give ShapeDWA the ability to average time series over the time domain. Second, the barycenter average in DBA is replaced by

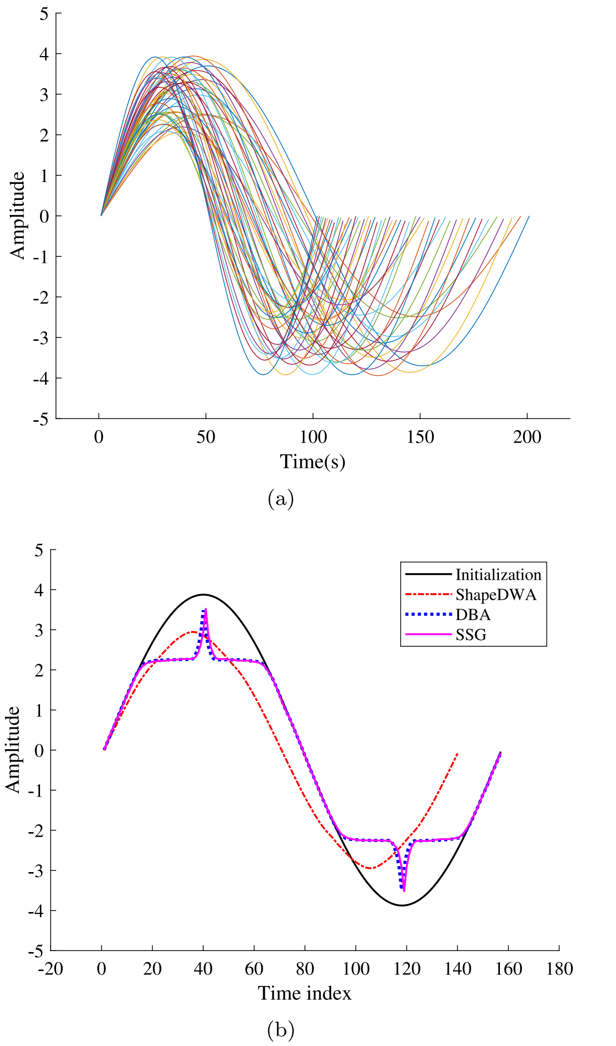

Fig. 2. Comparison of the DBA, SSG and ShapeDWA methods over a simulated dataset. (a) A set of sinusoidal signals with different periods and amplitudes. (b) Average time series calculated by DBA, SSG and ShapeDWA, respectively.

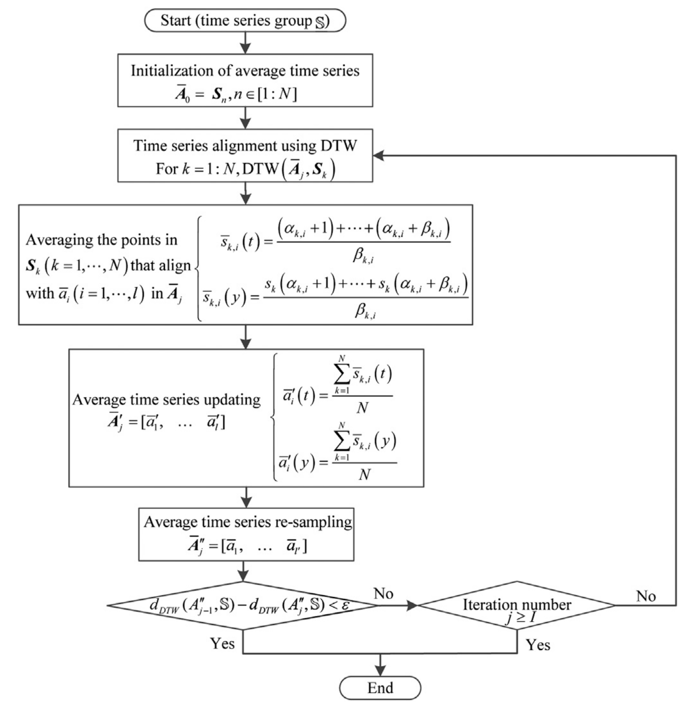

Fig. 3. The flowchart of the ShapeDWA algorithm.

a weighted average to eliminate the effect of noise and singular time series. The temporary average time series optimized by step (2) is a non-uniformly sampled average time series. Therefore, step (3) is added to re-sample the optimized average time series to obtain a uniformly sampled average time series. The flowchart of the ShapeDWA algorithm is shown in Fig. 3.

3.1. DTW weighted averaging optimization

Suppose $\bar{\boldsymbol{A}}_j = [\bar{a}_1, \ldots, \bar{a}_l]$ is the temporary average time series at the jth iteration. Taking the ith point $\bar{a}_i, i \in [1, l]$ of $\bar{\boldsymbol{A}}_j$ as an example, let us illustrate the optimization of $\bar{\boldsymbol{A}}_j$. Suppose $s_{k,i} = [s_k(\alpha_{k,i}+1), \ldots, s_k(\alpha_{k,i}+\beta_{k,i})]$ are the points in the kth time series $\boldsymbol{S}_k, k \in [1, N]$ that align with $\bar{a}_i$, where $(\cdot)$ is the time index. First, the points in each time series that align with $\bar{a}_i$ are averaged in both the time and amplitude domains,

$$\begin{cases} \bar{s}_{k,i}(t) = \frac{(\alpha_{k,i}+1) + \cdots + (\alpha_{k,i}+\beta_{k,i})}{\beta_{k,i}} \\ \bar{s}_{k,i}(y) = \frac{s_k(\alpha_{k,i}+1) + \cdots + s_k(\alpha_{k,i}+\beta_{k,i})}{\beta_{k,i}} \end{cases} \tag{7}$$

where $\bar{s}_{k,i}(t)$ and $\bar{s}_{k,i}(y)$ are the average time index and amplitude of the points aligning with $\bar{a}_i$ in $\boldsymbol{S}_k$, respectively.

Then, $\bar{a}_i$ is updated by taking the mean of the average alignment point $\bar{s}_{k,i} = [\bar{s}_{k,i}(t), \bar{s}_{k,i}(y)], \forall k \in [1, N]$ over all time series:

$$\begin{cases} \bar{a}'_i(t) = \frac{\sum_{k=1}^{N} \bar{s}_{k,i}(t)}{N} \\ \bar{a}'_i(y) = \frac{\sum_{k=1}^{N} \bar{s}_{k,i}(y)}{N} \end{cases} \tag{8}$$

Combined with Eqs. (7) and (8), we find that the temporary average time series is actually optimized by a weighted average function. Suppose $w_k(i)$ is the weight of the time series $\boldsymbol{S}_k$ to optimize the point $\bar{a}_i$. The weight $w_k(i)$ is determined based on the alignment between $\boldsymbol{S}_k$ and $\bar{\boldsymbol{A}}_j$ at point $\bar{a}_i$. If the alignment is a matching alignment or a stretching alignment, the corresponding weight satisfies that $w_k(i) = \frac{1}{N}$. If the alignment is a compressing alignment, then $w_k(i)$ is $\frac{1}{N \cdot \beta_{k,i}}$.

Assume $\bar{\boldsymbol{A}}'_j = [\bar{a}'_1, \ldots, \bar{a}'_l]$ is the update of $\bar{\boldsymbol{A}}_j$, where $\bar{a}'_i = [\bar{a}'_i(t), \bar{a}'_i(y)]$. It should be noted that this optimization does not change the length of the average time series. According to the step size constraint of DTW, the time index of the points in $\boldsymbol{S}_k$ that align with $\bar{a}'_{i+1}$ is greater than or equal to that of $\bar{a}'_i$. From Eqs. (7) and (8), we deduce that the time index of $\bar{\boldsymbol{A}}'_j$ is monotonically non-decreasing, i.e., $\bar{a}'_{i+1}(t) \geq \bar{a}'_i(t)$. The equal sign holds if and only if the alignments between $\bar{\boldsymbol{A}}$ and each time series in $\mathbb{S}$ at points $\bar{a}_i$ and $\bar{a}_{i+1}$ belong to the same group of stretching alignment. In this particular case, the amplitudes of the two points are also exactly the same, resulting in two consecutive identical points in $\bar{\boldsymbol{A}}'_j$.

3.2. Re-sampling

Due to the fact that time distortions inevitably exist between time series, the time index of the optimized average time series may not be all integers, resulting in a non-uniformly sampled time series. Therefore, the optimized average time series $\bar{\boldsymbol{A}}'_j$ should be re-sampled to a uniformly sampled time series.

Interpolation is a commonly used method for re-sampling the non-uniformly sampled time series. In this work, the Modified Akima Cubic Hermite interpolation (Makima) is utilized as it is parameterless and outperforms other interpolation techniques (MathWorks, 2019). Makima is a shape-preserving interpolation algorithm. Compared to the original Akima algorithm, Makima modifies the weighting strategy to give more weight to the side where the slope is closer to zero (Akima, 1970, 1974). As a result, it cannot only capture the movement between points to ensure the smoothness when the signal is oscillatory, but also avoid overshoots when the signal has flat areas.

The Makima interpolation method requires the timestamp of the time series to be unique. However, as mentioned above, there may be multiple consecutive repeat points in the optimized average time series $\bar{\boldsymbol{A}}'_j$. Therefore, the duplicate points in $\bar{\boldsymbol{A}}'_j$ should be eliminated before interpolation. Assume that $\bar{\boldsymbol{A}}''_j$ is the average time series after re-sampling. The length l'' of $\bar{\boldsymbol{A}}''_j$ depends on the time index of the last point $\bar{a}'_l$ of $\bar{\boldsymbol{A}}'_j$, i.e., $l'' = \mathrm{round}\left(\bar{a}'_l(t)\right)$, where $\mathrm{round}(\cdot)$ is the rounding function. According to the boundary constraint of DTW, the average time index of points in the kth time series $\boldsymbol{S}_k$ that align with $\bar{a}'_l$ is not greater than its length, i.e., $\bar{s}_{k,l}(t) \leq l_k$. As a result, the length l'' of $\bar{\boldsymbol{A}}''_j$ is not greater than the average length of time series in the dataset, i.e., $l'' \leq \sum_{k=1}^{N} \frac{l_k}{N}$. For each iteration, the length of the average time series varies, as shown in Fig. 4.

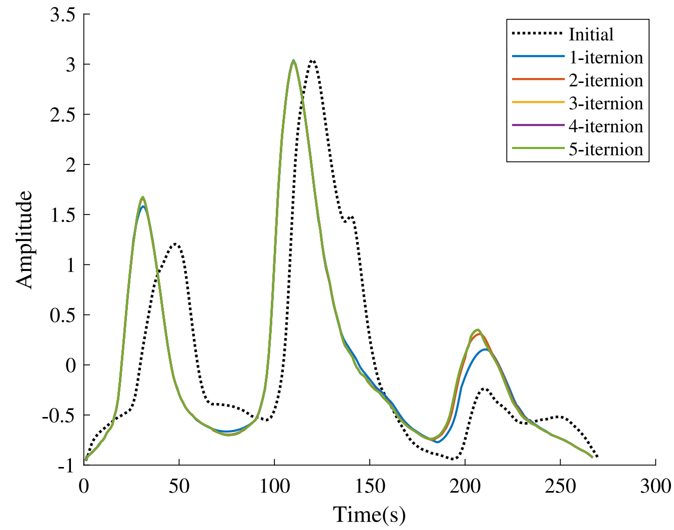

Fig. 4. The update of the average time series of the first class of the 50words dataset during the iterative optimization.

3.3. Convergence and computational complexity

Unlike DBA, ShapeDWA is not a monotonically convergent algorithm due to the process of weighted averaging and re-sampling. During optimization, the value of the average DTW distance may increase, but becomes stable after a finite number of iterations. Therefore, it is necessary to choose an appropriate termination criterion to ensure the quality and efficiency of the optimization process. In this work, two criteria are merged as the termination criterion of the iterative procedure: (T1) termination when a maximum number of iterations I has been exceeded; and (T2) termination when the fluctuation of the average DTW distance is less than a given threshold ε. Finally, the temporary average time series with the smallest average DTW distance is selected as the resulting average time series.

Fig. 4 illustrates the update of the average time series for the first class of the 50Words dataset during the iterative optimization process. The average time series tends to be stable after 3 iterations, exhibiting only minor variations near the extremes. These variations mainly due to the interpolation applied in the re-sampling process. Furthermore, the iterative optimization of the average time series verifies the capability of ShapeDWA to average the time series over the time domain. Fig. 5 displays the convergence curve of the average DTW distance computed by ShapeDWA with 10 different initializations. As observed from Fig. 5, the average DTW distance plateaus after 3 iterations, coinciding with the update of the average time series. The average DTW distance fluctuates around 6.15 with an amplitude of no more than 0.046. Although ShapeDWA is not a monotonically convergent algorithm, the iterative process does not diverge, and a satisfactory result can be obtained after a certain number of iterations. Generally, a higher number of iterations leads to a batter results. For this paper, the maximum number of iterations is set to 10, making a balance between computational complexity and accuracy. In addition, the termination threshold is set to 0.5% of the average DTW distance between the initial average time series and the time series set.

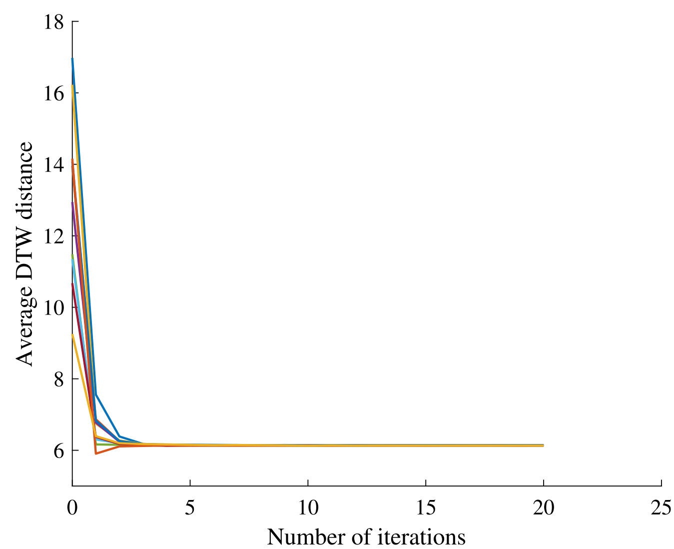

Fig. 5. The convergence curve of ShapeDWA on the first class of the 50words dataset with 10 different initializations.

Next, the complexity of ShapeDWA is detailed. As mentioned above, each iteration process can be divided into three steps. Since the complexity of DTW between $\bar{\boldsymbol{A}}_j$ and the kth time series $\boldsymbol{S}_k$ is $\Theta\left(l \cdot l_k\right)$, the time complexity of step (1) is $\Theta\left(l \cdot \sum_{k=1}^{N} l_k\right)$. In step (2), $\bar{\boldsymbol{A}}'_j$ is updated by taking the weighted average of all aligning points in both the amplitude and time domains. Consequently, step (2) has a time complexity of $\Theta\left(\sum_{k=1}^{N} l_k\right)$. Since the time complexity of the re-sampling process depends on the number of interpolation points, the maximum time complexity of step (3) is $\Theta\left(\sum_{k=1}^{N} \frac{l_k}{N}\right)$. Assuming that the iterative process is performed I times, the total time complexity of ShapeDWA is:

$$\Theta_{ShapeDWA} \leq I \cdot \left[\Theta\left(l \cdot \sum_{k=1}^{N} l_k\right) + \Theta\left(\sum_{k=1}^{N} l_k\right) + \Theta\left(\sum_{k=1}^{N} \frac{l_k}{N}\right)\right] \tag{9}$$

Considering $l \leq \sum_{k=1}^{N} \frac{l_k}{N}$, Eq. (9) can be simplified as:

$$\Theta_{ShapeDWA} \leq I \cdot \left[\frac{1}{N}\left(\sum_{k=1}^{N} l_k\right)^2 + \frac{N+1}{N} \sum_{k=1}^{N} l_k\right] \tag{10}$$

4. Metrics for evaluating the performance of DTW-based time series averaging algorithms

Evaluating the performance of a DTW-based time series averaging algorithm is a challenging task, as the ground truth average time series of multiple time series is unavailable. In previous work, the average DTW distance has been commonly used to evaluate time series averaging algorithms (Petitjean et al., 2011; Sun et al., 2016; Schultz

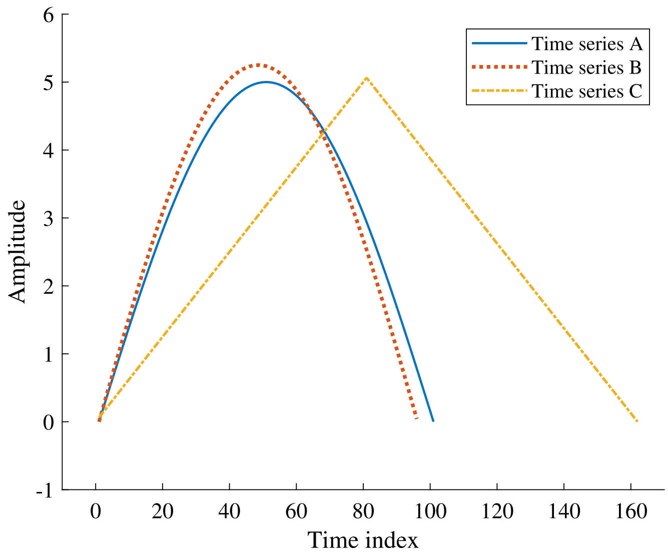

Fig. 6. Three time series for similarity evaluation.

and Jain, 2018). However, the DTW distance only takes into account the difference in the amplitude of time series. Therefore, this evaluation criterion may lead to a counterfactual result.

An ideal time series averaging algorithm should be able to average a set of time series in both the amplitude and time domains, while preserving the primary spatio-temporal characteristics of the time series dataset. In order to obtain an objective and comprehensive assessment, a suitable metric is required that is able to evaluate the average time series in terms of spatio-temporal characteristics. In this work, two metrics, the average discrepancy distance and the average time distortion, are introduced to evaluate the quality of the average time series.

4.1. Average discrepancy distance

To illustrate the limitations of using the DTW distance to evaluate the similarity of time series, let us begin by considering an example. Fig. 6 shows three time series. It is evident that time series A and time series B exhibit greater similarity. However, we obtain a wrong conclusion ($d_{DTW}(A,B) = 0.77$, $d_{DTW}(A,C) = 0.23$, $d_{DTW}(B,C) = 0.58$, time series A and time series C are the most similar based on DTW distance) when using the DTW distance as the similarity measure between time series. This is mainly due to the fact that the DTW distance only considers the differences in the amplitude of time series, but ignores the spatio-temporal characteristics of time series.

The first derivative of time series is a kind of spatio-temporal information that is appropriate as a local distance to measure the spatio-temporal differences between time series. Based on the sequence alignment calculated by DTW, a local distance that combines the differences of amplitude and the first derivative is developed to determine the discrepancy distance between two time series.

$$d_{st}(\boldsymbol{P},\boldsymbol{Q}) = \sum_{k=1}^{L}\left[w\left(p_{w_{k,1}} - q_{w_{k,2}}\right)^2 + (1-w)\left(\dot{p}_{w_{k,1}} - \dot{q}_{w_{k,2}}\right)^2\right] \tag{11}$$

where $w \in [0,1]$ is a positive weight, $\dot{p}_{w_{k,1}}$ and $\dot{q}_{w_{k,2}}$ are the first derivative of $\boldsymbol{P}$ and $\boldsymbol{Q}$, respectively. w is used to adjust the weight of the amplitude and the first derivative in the discrepancy distance. The modified discrepancy distance is able to evaluate the similarity of time series from the perspective of spatio-temporal. Assuming that the weights of the amplitude and the first derivative are equal, i.e., $w = 0.5$, the discrepancy distance between any two time series are $d_{st}(A,B) = 0.41$, $d_{st}(A,C) = 0.52$ and $d_{st}(B,C) = 0.90$, respectively. Obviously, the discrepancy distance obtains convincing results. The discrepancy distance is able to measure differences between time series in both the amplitude and time domains. Therefore, it is suitable for applications in which spatio-temporal characteristics predominate, such as gesture recognition, human motion assessment, handwriting recognition, path planning and evaluation.

Based on the discrepancy distance, we introduce a metric, the average discrepancy distance, to evaluate the quality of the average time series. It is defined as the mean of the discrepancy distances between the average time series and all time series,

$$D_{ad}\left(\bar{\boldsymbol{A}},\mathbb{S}\right) = \frac{1}{N}\sum_{k=1}^{N} d_{st}\left(\bar{\boldsymbol{A}},\boldsymbol{S}_k\right) \tag{12}$$

4.2. Average time distortion

As mentioned above, the time distortion measure provides a way to evaluate the similarity between two time series in the time domain. The time distortions between any two time series are $\rho(A,B) = 39$, $\rho(A,C) = 81$ and $\rho(B,C) = 100$, respectively. Similarly, the average time distortion measure is introduced to evaluate the time distortions between the average time series and all original time series, which is defined as

$$D_t\left(\bar{\boldsymbol{A}},\mathbb{S}\right) = \frac{1}{N}\sum_{j=1}^{N} \rho\left(\bar{\boldsymbol{A}},\boldsymbol{S}_j\right) \tag{13}$$

The average time distortion can be employed to qualify the ability of a time series averaging algorithm to average time series in the time domain. This metric is suitable for applications where timing or sequential order plays a critical role, such as gene sequencing, chemical reaction process control, medication analysis. A special type of time series, binary time series, should also be mentioned here. Because the temporal differences between binary time series are dominant, traditional amplitude-based metrics fail to quantify the differences between them. This is where the average time distortion comes in handy, providing for a reliable evaluation.

5. Experiment

In this section, we will conduct a series of experiments to evaluate the performance of the proposed ShapeDWA algorithm compared to the DBA and SSG methods.

5.1. Database and procedure

In order to obtain sufficient experimental results, both simulated and open-source time series datasets are utilized. The UCR Time Series Classification Archive database is a commonly used open source database for time series analysis (Dau et al., 2018). The database contains numerous types of data, such as images, motions, ECG, sensors, devices, spectro and simulated data. The number of classes varies between 2 and 60. According to the length of the time series in the dataset, the database can be divided into variable-length dataset and fixed-length dataset. Table 2 summarizes the selected time series.

The evaluation of DTW-based time series averaging algorithms is a non-trivial problem due to the lack of ground truth average time series as a Ref. Petitjean et al. (2011). In this paper, two experiments are performed to demonstrate the superiority of the proposed ShapeDWA algorithm against the DBA and SSG methods. In the first experiment, we systematically evaluate the performance of the proposed ShapeDWA algorithm, both visually and statistically. First, we perform a qualitative analysis by visually comparing the spatio-temporal characteristics of average time series. Since the DTW-based time series averaging is formulated as an optimization problem, the intra-class average discrepancy distance and intra-class average time distortion are then used for quantitative analysis. Finally, the effects of weighted averaging and initialization are discussed. The performance of the template matching-based classifier is highly dependent on the quality of the templates. In the second experiment, the template matching-based classification is utilized to evaluated these averaging methods, where the average time series of each class is used as the template.

In order to make these experiments reproducible, we provide details of the experimental settings here:

- For the DTW method, the local distance is the squared Euclidean distance.
- For each dataset, the training data and the test data are all merged. And we computed an average for each class in each dataset.
- The maximum number of iterations used for ShapeDWA, DBA, and SSG is 10, and the convergence threshold of ShapeDWA is set to 0.5% of the average DTW distance between the initial average time series and the time series set.
- The weight in the discrepancy distance is set to 0.5.

5.2. General performance comparison

5.2.1. Visual comparison

First, we show an average example of a simulated dataset. Fig. 2(a) illustrates a set of sinusoidal signals with different periods and amplitudes. The amplitude of these sinusoidal signals varies from 2 to 4, and the period varies from 1 s to 2 s with the step of 0.02 s. Since the sampling period is 0.01 s, the length of these sinusoidal signals varies from 100 to 200. Taking a randomly selected signal (its amplitude and period are 3.49 and 1.58 s) as the initial average time series, the average time series calculated by ShapeDWA, DBA and SSG are exhibited in Fig. 2(b). An ideal average time series should be a sinusoidal signal with average amplitude and period. As shown in Fig. 2(b), the waveform of the average time series calculated by ShapeDWA is close to a sinusoid, and its amplitude and period are 3.05 and 1.4 s, respectively. Obviously, it is more in line with our expectations. In contrast to the average time series calculated by DBA, it has two disadvantages. First, its shape is very different from the original signals, especially around the extreme points. Second, the length and the spatio-temporal characteristics of the average time series depend heavily on the choice of the initial time series. SSG faces the same problems. These demonstrate that the DBA and SSG algorithms are not capable of averaging time series in the time domain. This disadvantage will be more pronounced for the time series of unequal length and with large time distortions.

We then compare these three averaging methods on the real-world dataset from UCR database. Fig. 7(a) displays a set of time series in the first class of the 50words (image data) dataset, which are of equal length. Fig. 8(a) illustrates a set of time series in a class of the GestureMidAirD1 (motion data) dataset, which are of variable length. For these two datasets, the average time series calculated by these three methods are illustrated in Figs. 7(b) and 8(b), respectively. Similarly, visual comparisons indicate that the proposed ShapeDWA algorithm produces a superior average time series compared to the DBA and SSG algorithms.

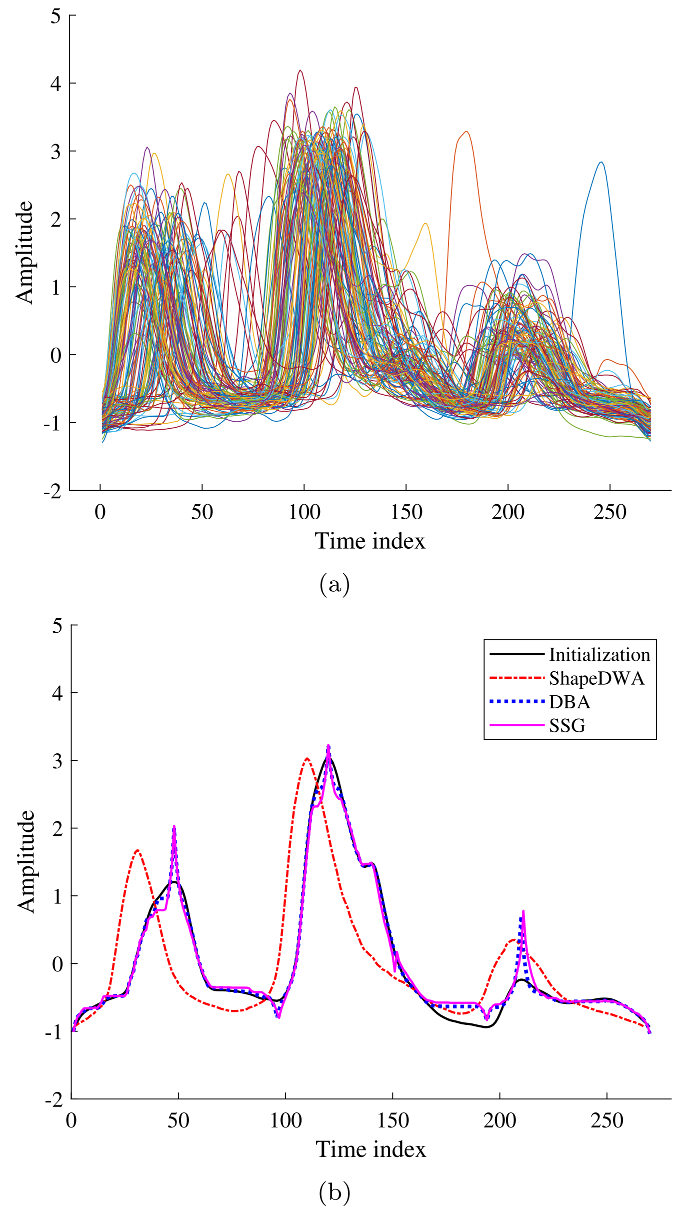

Fig. 7. Comparison of the DBA, SSG and ShapeDWA over the first class the 50words dataset (fixed-length). (a) The time series in the first class the 50words dataset. (b) Average time series calculated by DBA, SSG, and ShapeDWA, respectively.

5.2.2. Quantitative analysis

In this session, we will quantitatively compare ShapeDWA to DBA and SSG in terms of the intra-class average discrepancy distance and intra-class average time distortion. For these two metrics, the smaller the value, the better the average result. The average discrepancy distance of the simulated sinusoidal signals calculated by DBA, SSG, and ShapeDWA of 6.04, 5.27, and 5.04, and the average time distortion of 133.4, 128.9 and 83.1, respectively. For the 50words dataset, the average discrepancy distances calculated by DBA, SSG, and ShapeDWA are 6.31,6.06, and 3.68, and the average time distortions are 293.1, 291.0, and 220.3, respectively. The average discrepancy distances of the GestureMidAirD1 dataset calculated by DBA, SSG, and ShapeDWA are 1.811E+05, 2.073E+5, and 1.804E+05, and the average time distortions are 478.8, 479,2, and 472.8, respectively. These quantitative results are consistent with the previous visual comparison.

Tables 3 and 4 show the comparison of average discrepancy distance and average time distortion over the selected UCR datasets, respectively. Table 3 illustrates that ShapeDWA achieves a smaller average discrepancy distance than DBA in almost all of the datasets, except for the Haptics, Lighting2 and MoteStrain datasets. ShapeDWA outperforms DBA and SSG with an average reduction rate of 23.42% and 24.89%, respectively. A closer look at these three datasets reveals that there are some impulse noises of significant amplitude in these datasets. From Table 4, we notice that ShapeDWA reduces the average time distortion for all datasets compared with DBA and SSG, with an average reduction of 18.76% and 19.81%, respectively.

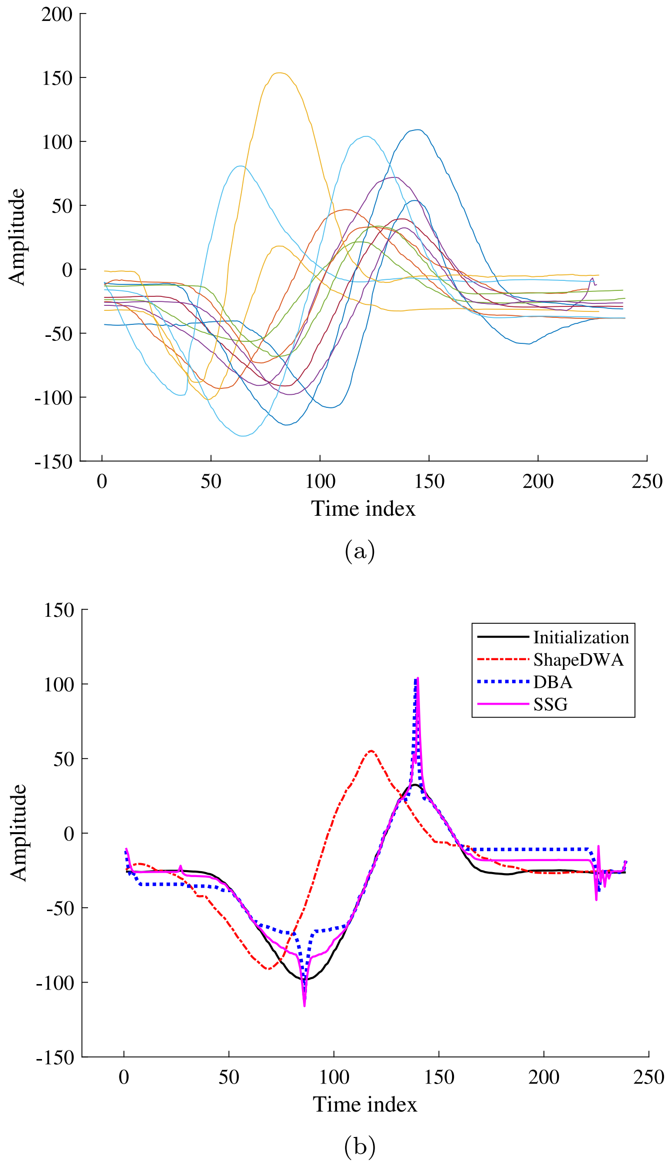

Fig. 8. Comparison of the DBA, SSG, and ShapeDWA methods over one class the GestureMidAirD1 dataset (variable-length). (a) The time series in one class the GestureMidAirD1. (b) Average time series calculated by DBA, SSG, and ShapeDWA, respectively.

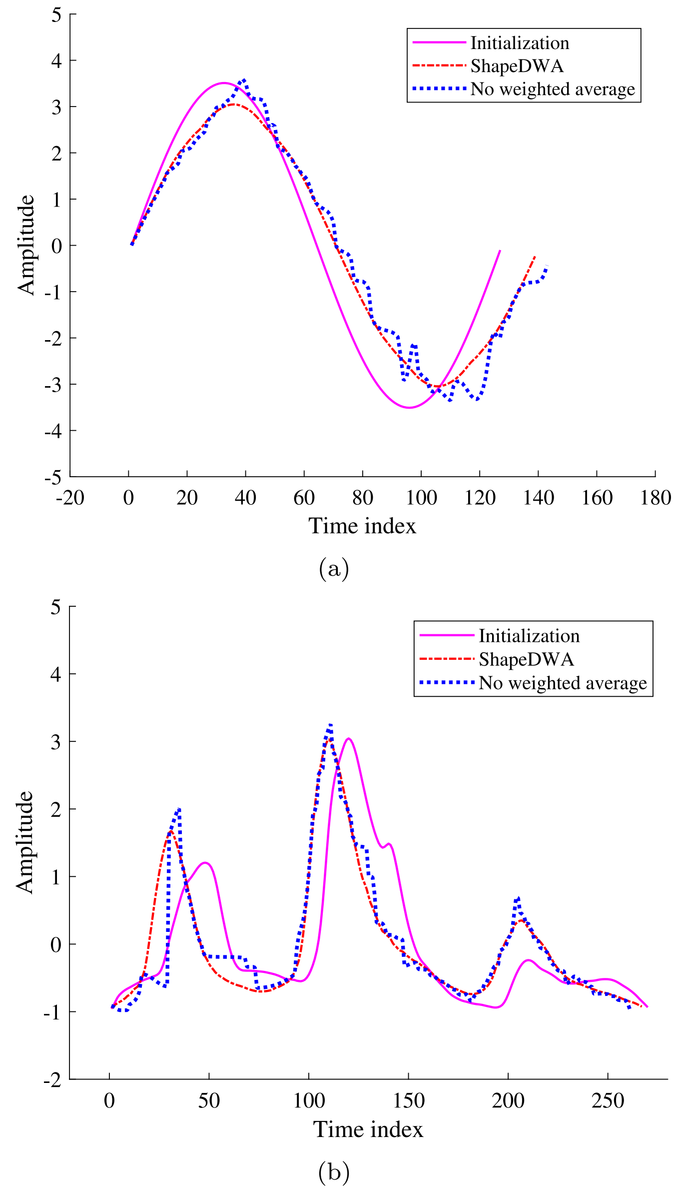

Fig. 9. Comparison of ShapeDWA with and without the weighted averaging. (a) Comparison over the sinusoidal signals. (b) Comparison over the 50words datasets.

5.2.3. *Impact of weighted averaging*

Fig. 9 displays the comparisons of ShapeDWA with and without the weighted averaging process on the sinusoidal signals and 50words datasets. It can be found from Fig. 9 that the average time series calculated by ShapeDWA with the weighted averaging process is smoother and preserves the shape characteristics of the original time series. On the contrary, the average time series computed by ShapeDWA without the weighted averaging process has many fluctuations that the original time series does not have. Obviously, such an average result most definitely well not have a positive impact on the further analysis. Further investigation shows that these fluctuations are primarily caused by the DTW pathological alignments due to noise, outliers and local amplitude differences. As a result, the weighted averaging process can effectively attenuate the impacts of the pathological alignment.

5.2.4. *Impact of initialization*

Similar to DBA, ShapeDWA starts with an initial average time series. A excellent algorithm should be robust to initialization. In this experiment, we will evaluate the effect of initialization by randomly selecting different individual time series from the dataset as the initial average time series.

Taking the simulated sinusoidal signals and the 50words datasets as examples, each time series in the dataset is selected as the initial average time series. With these initializations, the average time series calculated by DBA and ShapeDWA are shown in Figs. 10 and 11, respectively. Regardless of how the initial average time series is chosen, the average time series calculated by ShapeDWA are almost completely consistent. This confirms that ShapeDWA is capable of averaging time series in both the amplitude and time domains, and is robust to initialization, for both the fixed-length and variable-length time series. In contrary, the spatio-temporal characteristics (waveform) of average time series calculated by DBA are highly dependent on the initial average time series. This means that DBA is sensitive to initialization, especially for variable-length time series dataset.

To further explore the effect of initialization on ShapeDWA and DBA, we calculate the average discrepancy distance and average time distortion under different initializations. For the average time series

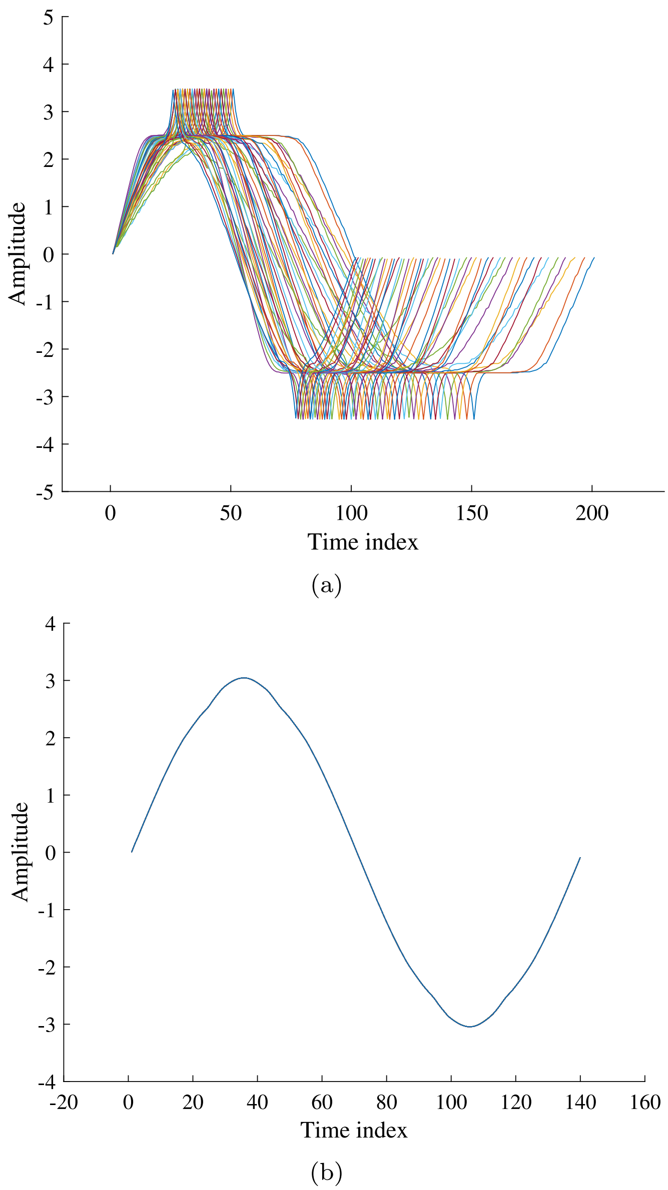

Fig. 10. Comparison of average time series calculated by DBA and ShapeDWA over the sinusoidal signals with different initializations. (a) Average time series calculated by DBA. (b) Average time series calculated by ShapeDWA.

shown in Figs. 10 and 11, the comparisons of average discrepancy distance and average time distortion are presented in Figs. 12 and 13, respectively. Due to the limitation of computational constraints, we selected a subset of the dataset to further evaluate the robustness of ShapeDWA to initialization. For each class in each dataset, 50% of the time series are randomly selected as the initial average time series. Figs. 14 and 15 presents the mean and standard deviation of the average discrepancy distance and average time distortion on a few datasets. Since the amplitude and length of each individual dataset differ by orders of magnitude, these data are normalized for exhibition. With different initializations, the average discrepancy distance and average time distortion calculated by ShapeDWA are not only smaller but also more stable than those calculated by DBA for all of the datasets. These comparison results demonstrate that ShapeDWA is more robust to initialization than DBA is.

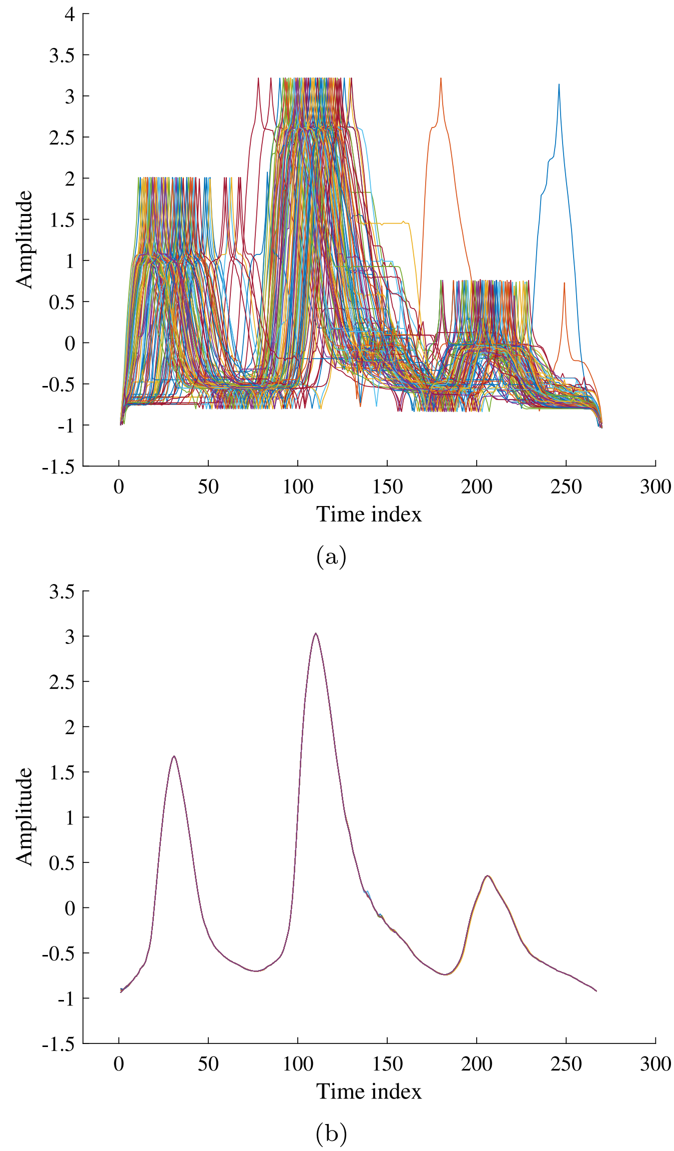

Fig. 11. Comparison of average time series calculated by DBA and ShapeDWA over the 50words datasets with different initializations. (a) Average time series calculated by DBA. (b) Average time series calculated by ShapeDWA.

5.3. Template matching based classification

In this experiment, we evaluate the performance of the proposed ShapeDWA algorithm through template matching based time series classification over some UCR datasets. A higher classification rate indicates that the average algorithm can produce a template that represents the majority of the features of the intra-class data. To make the experimental results more convincing, the 5-fold cross-validation procedure is employed. For each dataset, 50% of the time series are randomly selected as the training dataset and the rest as the test dataset. For each class in each dataset, one time series is randomly selected from the training dataset as the initial time series to calculate the average time series. Three types of templates are utilized for classification, including initial time series, average time series calculated by DBA and ShapeDWA.

Using the DTW distance, the discrepancy distance (w = 0.5), and the time distortion measure as the similarity metric, Table 5 shows the comparisons of the classification rate based on different templates. From Table 5 we can see that ShapeDWA outperforms DBA (SSG) for

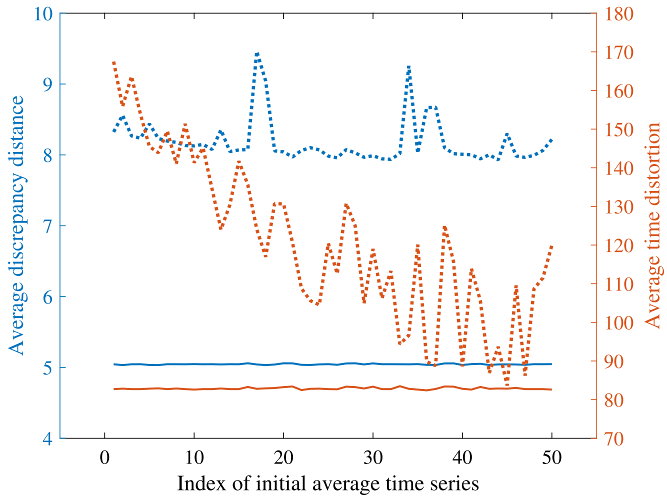

Fig. 12. Comparison of average discrepancy distance and average time distortion calculated by ShapeDWA (solid line) and DBA (dashed line) with different initializations over the simulated sinusoidal signals.

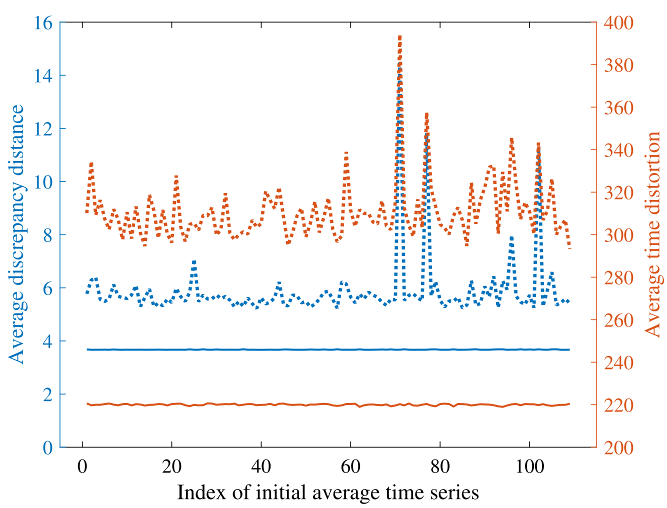

Fig. 13. Comparison of average discrepancy distance and average time distortion calculated by ShapeDWA (solid line) and DBA (dashed line) with different initializations over the 50words datasets.

most of the datasets, with an average improvement of 5.05% (−1.17%), 14.3%(23.26%) and 31.86%(27.16%) for DTW distance, discrepancy distance and time distortion measure, respectively. Utilizing DTW distance and discrepancy distance as the similarity metrics, templates calculated by these three time series averaging methods always outperform randomly selected templates. Obviously, the average time series is more representative than any individual instance. Using the average time series as templates, the classification rate of the template matching based classifier can be significantly improved, even comparable to the K-nearest neighbor classifier. It is suitable for a system with limited resources in either memory storage or computational power, e.g., micro-sensors, wearable devices. For the DBA and SSG algorithms, using DTW distance as a similarity metric always achieves a higher classification rate than using the discrepancy distance and the time distortion measure. Using time distortion measure as a similarity metric, DBA and SSG obtain even worse results than randomly selecting templates. This is mainly due to the fact that DBA and SSG lack the

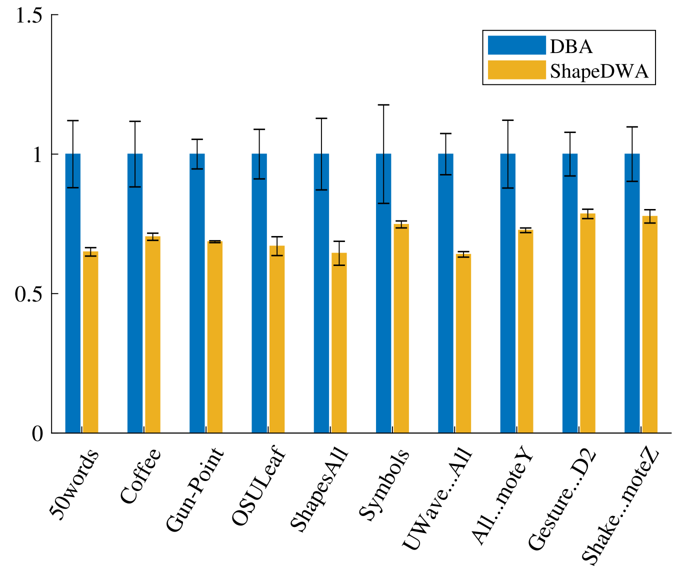

Fig. 14. Comparison of average discrepancy distance obtained by DBA and ShapeDWA with different initializations.

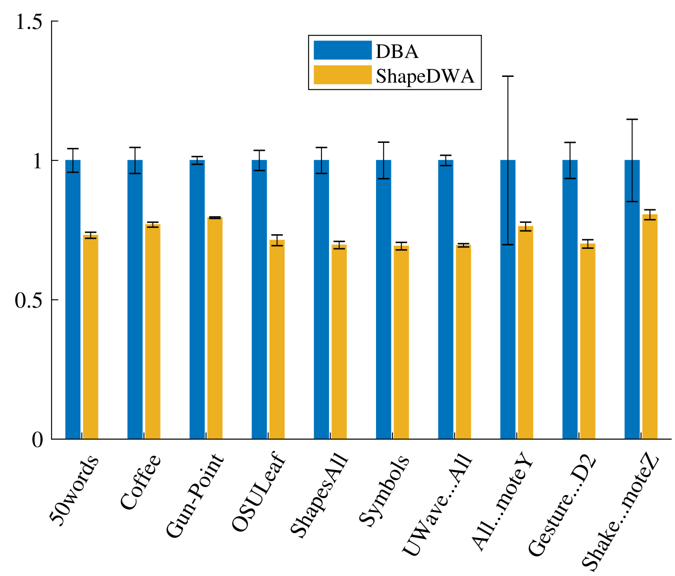

Fig. 15. Comparison of average time distortion obtained by DBA and ShapeDWA with different initializations.

ability to average the time series in the time domain, resulting in a large time distortion in the averaged time series.

5.4. Discussion

In summary, the above experimental results demonstrate that the proposed ShapeDWA algorithm is superior than the DBA and SSG algorithms. First, the visual comparison experiments show that ShapeDWA is a shape-preserving averaging algorithm. This means that ShapeDWA can maintain the spatio-temporal characteristics of the time series dataset. Next, a series of quantitative analysis experiments demonstrate that ShapeDWA has the ability to average a set of time series in both the amplitude and time domains, solving the problem of DBA being sensitive to initialization. Finally, the combination of the ShapeDWA algorithm with the two similarity measures can significantly improve the classification rate of the template-based matching classifier.

6. Conclusion and further work

In this paper, we have proposed a novel shape-based time series averaging algorithm that is capable of averaging time series in both the amplitude and time domains. Although the algorithm was initially designed for univariate time series, it can be extended to multivariate time series by replacing the local distance, averaging, and interpolation functions. The experimental results on the simulated and open source datasets demonstrate that the proposed ShapeDWA algorithm outperforms the DBA and SSG algorithms both visually and statistically. By utilizing the algorithm, a reference template can be created for applications such as motion evaluation, motion planning, gesture recognition, and signature recognition.

Although the experimental results show the superior performance of ShapeDWA, it still has the potential for improvement. In future work, we will improve ShapeDWA in the following directions. Firstly, the original DTW algorithm used for time series alignment will be replaced by some proper DTW variants that can capture spatio-temporal similarity between time series. Time series alignment is the basis of ShapeDWA, which directly affects the performance of optimization. However, the original DTW algorithm suffers from the pathological alignment problem due to noise and local amplitude, resulting in unreasonable and unintuitive alignment (Keogh and Pazzani, 2001). For instance, a local peak aligns with a local valley. Obviously, the average optimization based on this pathological alignment cannot generate an average time series that preserves the spatio-temporal characteristics of the dataset. Therefore, several DTW variants have been developed to overcome this problem by adding warping path constraint, local weight, or local feature distance, including derivative DTW (DDTW) (Keogh and Pazzani, 2001), EventDTW (Jiang et al., 2020), ShapeDTW (Zhao and Itti, 2018; Belongie et al., 2002), constrained DTW (Itakura, 1975; Li et al., 2020; Morel et al., 2018) and weighted DTW (Maus et al., 2016; Jeong et al., 2011). For instance, a modified DBA that utilizes a constrained DTW can effectively reduce the singularities in the averaged series in the averaged time series (Morel et al., 2018). Second, the re-sampling and iterative termination strategy needs further optimization to improve computational efficiency. As mentioned above, ShapeDWA is not monotonically convergent due to the re-sample process. Therefore, some proper restrictions need to be added to the re-sample process to avoid oscillations in the optimization process. I believe that this research can open up a wide range of perspectives in the time series averaging under DTW.

CRediT authorship contribution statement

Yutao Liu: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Visualization. **Yong-An Zhang:** Writing – review & editing, Validation. **Ming Zeng:** Writing – review & editing. **Jie Zhao:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was partially supported by the Joint Construction Project for Medical Science and Technology of Henan Province, China under Grant (LHGJ20210341).

Appendix

See Tables 2–5.

Table 2
Summarization of the selected time series in the UCR time series classification archive database.

Dataset	Type	Class	Num	Length	Dataset	Type	Class	Num	Length
50words	Image	50	905	270	Non...ax1	ECG	42	3765	750
Adiac	Image	37	781	176	Non...ax2	ECG	42	3765	750
ArrowHead	Image	3	211	251	OSULeaf	Image	6	442	427
Beef	Spectro	5	60	470	OliveOil	Spectro	4	60	570
BeetleFly	Image	2	40	512	Pha...ect	Image	2	2658	80
BirdChicken	Image	2	40	512	Phoneme	Sensor	39	2110	1024
CBF	Simulated	3	930	128	Plane	Sensor	7	210	144
Car	Sensor	4	120	577	Pro...oup	Image	3	605	80
Chl...ion	Sensor	3	4307	166	Pro...ect	Image	2	891	80
Cin...rso	Sensor	4	1420	1639	Pro...xTW	Image	6	605	80
Coffee	Spectro	2	56	286	Ref...ces	Device	3	750	720
Computers	Device	2	500	720	ScreenType	Device	3	750	720
Cricket_X	Motion	12	780	300	ShapeletSim	Simulated	2	200	500
Cricket_Y	Motion	12	780	300	ShapesAll	Image	60	1200	512
Cricket_Z	Motion	12	780	300	Sma...ces	Device	3	750	720
Dia...ion	Image	4	322	345	Son...ace	Sensor	2	621	70
Dis...roup	Image	3	539	80	Son...eII	Sensor	2	980	65
Dis...ect	Image	2	876	80	Sta...ves	Sensor	3	9236	1024
Dis...xTW	Image	6	539	80	Strawberry	Spectro	2	983	235
ECG200	ECG	2	200	96	SwedishLeaf	Image	15	1125	128
ECG5000	ECG	5	5000	140	Symbols	Image	6	1020	398
ECGFiveDays	ECG	2	884	136	Toe...on1	Motion	2	268	277
Earthquakes	Sensor	2	461	512	Toe...on2	Motion	2	166	343
ElectricDevices	Device	7	16 637	96	Trace	Sensor	4	200	275
FISH	Image	7	350	463	TwoLeadECG	ECG	2	1162	82
FaceAll	Image	14	2250	131	Two_Patterns	Simulated	4	5000	128
FaceFour	Image	4	112	350	UWa...All	Motion	8	4478	945
FacesUCR	Image	14	2250	131	Wine	Spectro	2	111	234
FordA	Sensor	2	4921	500	Wor...yms	Image	25	905	270

(*continued on next page*)

Table 2 (*continued*).

Dataset	Type	Class	Num	Length	Dataset	Type	Class	Num	Length
FordB	Sensor	2	4446	500	Worms	Motion	5	258	900
Gun_Point	Motion	2	200	150	Wor...ass	Motion	2	258	900
Ham	Spectro	2	214	431	syn...trol	Simulated	6	600	60
HandOutlines	Motion	2	1370	2709	uWa...ry_X	Motion	8	4478	315
Haptics	Motion	5	463	1092	uWa...ry_Y	Motion	8	4478	315
Herring	Image	2	128	512	uWa...ry_Z	Motion	8	4478	315
InlineSkate	Motion	7	650	1882	wafer	Sensor	2	7164	152
Ins...und	Sensor	11	2200	256	yoga	Spectro	2	3300	426
Ita...and	Sensor	2	1096	24	All...teX	Sensor	10	1000	Vary
Lar...ces	Device	3	750	720	All...teY	Sensor	10	1000	Vary
Lighting2	Sensor	2	121	637	All...teZ	Sensor	10	1000	Vary
Lighting7	Sensor	7	143	319	Ges...rD1	Trajectory	26	338	Vary
MALLAT	Simulated	8	2400	1024	Ges...rD2	Trajectory	26	338	Vary
Meat	Spectro	3	120	448	Ges...rD3	Trajectory	26	338	Vary
MedicalImages	Image	10	1141	99	Ges...eZ1	Sensor	6	304	Vary
Mid...oup	Image	3	554	80	Ges...eZ2	Sensor	6	304	Vary
Mid...ect	Image	2	891	80	Pic...teZ	Sensor	10	100	Vary
Mid...xTW	Image	6	553	80	PLAID	Device	11	1074	Vary
MoteStrain	Sensor	2	1272	84	Sha...teZ	Sensor	10	100	Vary

Table 3
Comparison of average discrepancy distance (lower is better) calculated by ShapeDWA, DBA, and SSG over the selected UCR datasets.

Dataset	ShapeDWA	DBA	SSG	Dataset	ShapeDWA	DBA	SSG
50words	6.85	10.95	12.07	Non...ax1	1.24	1.60	1.61
Adiac	0.17	0.21	0.21	Non...ax2	1.09	1.38	1.39
ArrowHead	2.77	4.02	3.91	OSULeaf	25.79	37.66	37.80
Beef	13.78	20.80	30.62	OliveOil	0.01	0.01	0.01
BeetleFly	28.06	46.24	45.18	Pha...ect	1.38	1.71	1.73
BirdChicken	23.76	32.71	29.87	Phoneme	449.25	548.71	556.58
CBF	43.36	49.56	49.07	Plane	0.49	0.63	0.61
Car	1.45	2.02	2.31	Pro...oup	0.36	0.39	0.40
Chl...ion	41.50	49.75	49.26	Pro...ect	0.58	0.70	0.71
Cin...rso	203.22	275.40	285.35	Pro...xTW	0.32	0.38	0.39
Coffee	0.70	0.95	0.96	Ref...ces	507.63	815.56	967.84
Computers	701.95	1124.40	1016.86	ScreenType	639.14	1055.27	1097.58
Cricket_X	84.67	109.60	113.51	ShapeletSim	603.39	619.19	634.53
Cricket_Y	61.97	84.79	88.50	ShapesAll	11.82	17.73	17.34
Cricket_Z	83.38	108.39	112.05	Sma...ces	2.17E+03	3.06E+03	3.03E+03
Dia...ion	1.77	2.02	3.42	Son...ace	11.22	14.05	12.84
Dis...roup	1.16	1.45	1.49	Son...eII	24.54	27.99	29.59
Dis...ect	2.16	2.83	2.83	Sta...ves	13.35	16.86	17.11
Dis...xTW	0.91	1.22	1.22	Strawberry	1.63	2.21	2.25
ECG200	12.52	15.30	14.96	SwedishLeaf	2.55	3.46	3.44
ECG5000	22.47	31.14	29.74	Symbols	2.84	3.62	3.94
ECGFiveDays	8.61	9.47	9.70	Toe...on1	41.55	64.10	64.27
Earthquakes	951.13	1308.36	1274.78	Toe...on2	53.92	88.29	94.92
ElectricDevices	163.96	200.79	233.27	Trace	1.66	2.21	2.11
FISH	7.31	7.36	7.40	TwoLeadECG	1.42	1.72	1.97
FaceAll	22.96	27.99	27.45	Two_Patterns	43.78	69.65	70.95
FaceFour	52.48	60.66	60.17	UWa...All	83.39	125.84	105.58
FacesUCR	22.86	26.09	26.45	Wine	0.06	0.07	0.07
FordA	63.81	90.73	96.53	Wor...yms	15.49	25.73	25.24
FordB	69.30	95.35	102.12	Worms	169.89	245.34	285.24
Gun_Point	4.09	6.72	6.26	Wor...ass	169.39	271.18	293.04
Ham	27.18	39.11	39.40	synthetic_control	29.56	31.01	30.98
HandOutlines	8.42	10.82	15.24	uWa...ry_X	31.74	48.56	44.78
Haptics	40.53	30.15	29.37	uWa...ry_Y	31.47	46.43	40.23
Herring	0.92	1.42	1.32	uWa...ry_Z	31.77	53.55	45.84
InlineSkate	88.06	134.23	101.38	wafer	78.27	136.68	150.54
Ins...und	14.29	25.24	24.37	yoga	17.61	30.57	22.12
Ita...and	2.93	3.24	3.35	All...teX	17.71	29.51	97.73
Lar...ces	659.24	861.28	734.61	All...teY	11.28	17.51	19.46
Lighting2	400.05	366.06	355.95	All...teZ	11.40	17.45	17.63
Lighting7	174.11	165.60	181.87	Ges...rD1	4.47E+04	5.57E+04	6.43E+04
MALLAT	4.47	4.91	4.69	Ges...rD2	6.06E+04	7.64E+04	9.07E+04
Meat	0.04	0.05	0.05	Ges...rD3	2.34E+05	2.87E+05	3.78E+05
MedicalImages	15.50	19.18	18.69	Ges...eZ1	4.03E+03	4.60E+03	4.53E+03
Mid...oup	0.58	0.65	0.65	Ges...eZ2	3.83E+03	4.68E+03	4.65E+03
Mid...ect	0.85	1.13	1.14	Pic...teZ	1.19E+05	1.93E+05	2.20E+05
Mid...xTW	0.55	0.65	0.64	PLAID	1.10	1.69	1.70
MoteStrain	48.34	46.78	47.10	Sha...teZ	9.02	11.31	11.69

Table 4
Comparison of average time distortion (lower is better) calculated by ShapeDWA, DBA, and SSG over the selected UCR datasets.

Dataset	ShapeDWA	DBA	SSG	Dataset	ShapeDWA	DBA	SSG
50words	228.90	289.48	271.56	Non...ax1	639.41	710.33	703.09
Adiac	27.18	34.19	33.91	Non...ax2	602.32	675.52	669.34
ArrowHead	167.10	209.70	212.06	OSULeaf	388.04	489.93	500.22
Beef	478.30	574.08	548.42	OliveOil	105.37	127.33	122.55
BeetleFly	407.50	533.23	490.69	Pha...ect	35.76	39.43	39.34
BirdChicken	442.98	619.95	559.55	Phoneme	1079.42	1231.88	1229.90
CBF	98.21	112.48	120.17	Plane	24.95	31.12	31.33
Car	324.42	386.75	385.91	Pro...oup	13.86	17.45	18.42
Chl...ion	95.51	116.69	126.44	Pro...ect	25.36	28.37	28.25
Cin...rso	2360.48	2815.44	2927.92	Pro...xTW	14.41	16.34	16.50
Coffee	166.25	201.25	196.78	Ref...ces	690.47	938.19	1009.82
Computers	743.80	926.41	981.62	ScreenType	758.06	983.70	1030.04
Cricket_X	364.53	437.25	441.29	ShapeletSim	300.73	341.31	352.97
Cricket_Y	354.33	421.06	428.16	ShapesAll	349.85	464.49	435.17
Cricket_Z	363.94	435.52	441.25	Sma...ces	493.13	612.01	636.03
Dia...ion	94.83	123.88	124.69	Son...ace	39.56	47.65	46.47
Dis...roup	23.58	28.51	28.53	Son...eII	36.63	43.11	47.68
Dis...ect	37.21	41.47	40.96	Sta...ves	921.32	1254.92	1407.71
Dis...xTW	20.55	24.98	25.21	Strawberry	188.07	205.88	211.46
ECG200	81.21	94.47	96.70	SwedishLeaf	61.97	71.36	71.14
ECG5000	149.08	174.98	178.38	Symbols	241.11	324.31	337.77
ECGFiveDays	126.49	150.76	158.98	Toe...on1	312.17	368.38	375.43
Earthquakes	326.22	427.95	431.79	Toe...on2	429.63	504.81	509.88
ElectricDevices	59.84	72.47	81.69	Trace	284.54	318.62	320.38
FISH	187.78	234.69	235.90	TwoLeadECG	42.99	50.22	53.77
FaceAll	65.28	79.19	81.97	Two_Patterns	78.48	107.13	114.16
FaceFour	213.50	269.70	275.36	UWa...All	982.70	1314.75	1383.14
FacesUCR	65.28	78.58	81.30	Wine	32.46	39.42	39.57
FordA	397.55	452.19	458.56	Wor...yms	279.28	343.13	334.23
FordB	405.10	451.80	469.87	Worms	1232.95	1453.17	1458.95
Gun_Point	175.32	209.15	208.00	Wor...ass	1252.99	1456.98	1516.87
Ham	432.77	492.45	501.51	synthetic_control	34.35	38.95	40.16
HandOutlines	1038.85	1647.00	2314.25	uWa...ry_X	315.26	438.64	455.73
Haptics	853.01	1145.34	1173.34	uWa...ry_Y	340.83	457.20	483.43
Herring	271.46	348.32	348.07	uWa...ry_Z	322.50	446.25	460.42
InlineSkate	2310.92	2939.87	3004.56	wafer	174.70	233.93	253.84
Ins...und	268.24	324.33	336.52	yoga	354.03	498.23	536.36
Ita...and	11.79	13.99	14.87	All...teX	146.78	184.57	186.89
Lar...ces	643.81	887.54	861.28	All...teY	151.60	196.00	201.31
Lighting2	816.79	1005.07	1004.13	All...teZ	144.31	174.74	178.65
Lighting7	322.71	408.41	395.59	Ges...rD1	166.05	228.33	203.51
MALLAT	694.18	864.87	878.28	Ges...rD2	177.01	241.63	215.42
Meat	131.57	164.14	162.01	Ges...rD3	173.08	257.40	243.29
MedicalImages	103.53	121.64	121.55	Ges...eZ1	234.47	271.21	270.60
Mid...oup	22.09	25.09	25.01	Ges...eZ2	233.31	269.85	270.38
Mid...ect	32.45	38.88	38.81	Pic...teZ	402.12	633.72	545.32
Mid...xTW	22.43	25.19	25.09	PLAID	125.67	148.93	145.21
MoteStrain	74.32	98.98	110.46	Sha...teZ	141.82	165.41	161.30

Table 5
Comparison of classification accuracy (higher is better) based on different templates and similarity measure.

Dataset	DTW distance				Discrepancy distance				Time distortion			
	Random	DBA	SSG	ShapeDWA	Random	DBA	SSG	ShapeDWA	Random	DBA	SSG	ShapeDWA
50words	41.30	50.67	53.24	57.45	43.90	51.35	48.64	63.50	44.76	35.42	48.73	**64.15**
Adiac	39.44	46.40	46.65	48.17	43.78	45.63	46.95	**53.10**	41.88	39.59	44.01	51.13
Car	41.67	53.33	61.67	58.67	42.87	47.67	31.33	58.87	50.67	49.67	56.00	**64.53**
CinC...rso	32.56	33.35	48.76	37.92	31.03	35.57	33.85	37.33	41.94	34.38	33.71	**52.94**
ECG200	59.80	69.50	69.90	73.07	51.19	68.29	68.32	**76.25**	60.70	68.08	71.49	72.67
FISH	45.11	52.80	64.57	57.71	45.83	48.23	45.37	62.40	53.03	53.14	53.71	**72.49**
FaceAll	56.12	82.03	85.07	82.82	57.68	78.07	76.19	**85.05**	63.56	70.69	67.77	84.01
FacesUCR	53.23	83.10	85.98	83.87	57.62	75.35	75.34	**84.91**	60.09	70.46	66.86	81.19
Gun_Point	57.80	57.60	65.00	57.40	57.40	61,83	57.40	68.20	60.80	62.60	64.80	**71.80**
Haptics	28.02	32.67	32.50	25.17	25.69	27.93	27.84	26.55	25.60	28.71	28.36	**40.30**
Lighting2	55.41	60.00	60.98	**60.20**	47.21	49.51	49.51	45.25	50.82	48.85	46.56	55.41
MoteStrain	67.18	81.30	85.37	84.27	59.72	67.69	69.70	74.03	64.11	69.20	69.89	**85.53**
Non...rax2	58.40	75.72	76.11	77.56	60.21	69.88	70.08	**82.26**	43.31	54.70	56.78	64.91
OSULeaf	31.02	30.15	45.68	36.77	28.20	27.12	25.32	37.31	37.48	34.68	29.82	**38.96**
Plane	97.33	99.62	99.62	100.00	97.33	99.05	98.67	99.62	98.86	99.62	99.62	100.00
ScreenType	35.31	40.48	39.95	40.59	34.51	35.79	35.09	**43.57**	34.51	37.49	36.05	39.36

(*continued on next page*)

Table 5 (*continued*).

Dataset	DTW distance				Discrepancy distance				Time distortion			
	Random	DBA	SSG	ShapeDWA	Random	DBA	SSG	ShapeDWA	Random	DBA	SSG	ShapeDWA
ShapesAll	48.17	55.43	58.70	59.17	48.20	50.23	45.43	59.97	51.23	40.20	50.30	**62.47**
Sony...face	62.65	93.89	94.02	94.02	62.30	83.92	85.72	**93.31**	64.82	73.61	77.49	87.56
SwedishLeaf	49.79	65.39	69.09	66.04	50.49	56.44	51.82	**75.44**	54.67	50.54	57.89	67.02
Symbols	88.05	92.27	97.66	94.69	88.36	86.64	82.58	**95.04**	82.30	67.81	67.07	88.98
Toe...tion1	54.63	62.69	62.09	69.82	59.12	64.63	60.15	70.60	61.94	69.70	70.60	**79.37**
Trace	78.60	92.60	93,41	93.84	78.00	86.00	85.60	**95.80**	60.40	53.20	52.40	60.60
Two_Patterns	91.53	97.30	97.72	**98.84**	54.41	58.55	59.89	85.89	46.93	54.02	53.89	84.39
UWave...All	49.97	63.26	73.57	**71.16**	44.70	49.64	47.74	68.13	45.68	27.02	28.48	61.96
Wine	49.14	59.00	59.29	59.50	51.43	58.50	58.31	64.29	54.71	55.36	55.71	**65.21**
wafer	51.81	55.81	60.63	**76.74**	54.73	60.21	46.16	61.58	62.57	60.36	53.54	54.72
All...moteX	32.96	44.12	46.56	48.88	33.88	47.32	29.64	**51.20**	19.84	17.88	18.36	27.56
Ges...eZ2	33.64	58.83	58.83	57.45	34.42	62.91	57.66	55.32	47.53	47.27	48.44	**64.69**
Pick...moteZ	54.80	65.40	65.60	66.80	53.20	62.40	54.80	**73.80**	28.80	24.80	28.00	34.00
Shake...imoteZ	65.60	78.80	74.00	80.80	67.60	80.80	75.20	**84.60**	36.40	30.80	34.00	52.20

References

Akima, H., 1970. A new method of interpolation and smooth curve fitting based on local procedures. J. ACM 17 (4), 589–602.

Akima, H., 1974. A method of bivariate interpolation and smooth surface fitting based on local procedures. Commun. ACM 17 (1), 18–20.

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sens. Environ. 204, 509–523.

Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 24 (4), 509–522.

Brill, M., Fluschnik, T., Froese, V., Jain, B., Niedermeier, R., Schultz, D., 2019. Exact mean computation in dynamic time warping spaces. Data Min. Knowl. Discov. 33 (1), 252–291.

Dau, H.A., Bagnall, A., Kamgar, K., Yeh, C.-C.M., Zhu, Y., Gharghabi, S., Ratanamahatana, C.A., Keogh, E., 2019. The UCR time series archive. IEEE/CAA J. Autom. Sin. 6 (6), 1293–1305.

Dau, H.A., Keogh, E., Kamgar, K., Yeh, C.-C.M., Zhu, Y., Gharghabi, S., Ratanamahatana, C.A., Yanping, Hu, B., Begum, N., Bagnall, A., Mueen, A., Batista, G., Hexagon-ML, 2018. The UCR time series classification archive. https://www.cs.ucr.edu/~eamonn/time_series_data_2018/.

D'Urso, P., De Giovanni, L., Massari, R., 2021. Trimmed fuzzy clustering of financial time series based on dynamic time warping. Ann. Oper. Res. 299 (1), 1379–1395.

Folgado, D., Barandas, M., Matias, R., Martins, R., Carvalho, M., Gamboa, H., 2018. Time alignment measurement for time series. Pattern Recognit. 81, 268–279.

Fu, A.W.-C., Keogh, E., Lau, L.Y.H., Ratanamahatana, C.A., Wong, R.C.-W., 2008. Scaling and time warping in time series querying. VLDB J. 17 (4), 899–921.

Gupta, L., Molfese, D.L., Tammana, R., Simos, P.G., 1996. Nonlinear alignment and averaging for estimating the evoked potential. IEEE Trans. Biomed. Eng. 43 (4), 348–356.

Gusfield, D., 1997. Algorithms on stings, trees, and sequences: Computer science and computational biology. Acm Sigact News 28 (4), 41–60.

Hachaj, T., Piekarczyk, M., Ogiela, M.R., 2017. Human actions analysis: Templates generation, matching and visualization applied to motion capture of highly-skilled karate athletes. Sensors 17 (11), 2590.

Itakura, F., 1975. Minimum prediction residual principle applied to speech recognition. IEEE Trans. Acoust. Speech Signal Process. 23 (1), 67–72.

Izakian, H., Pedrycz, W., Jamal, I., 2015. Fuzzy clustering of time series data using dynamic time warping distance. Eng. Appl. Artif. Intell. 39, 235–244.

Jeong, Y.-S., Jeong, M.K., Omitaomu, O.A., 2011. Weighted dynamic time warping for time series classification. Pattern Recognit. 44 (9), 2231–2240.

Jiang, Y., Qi, Y., Wang, W.K., Bent, B., Avram, R., Olgin, J., Dunn, J., 2020. EventDTW: An improved dynamic time warping algorithm for aligning biomedical signals of nonuniform sampling frequencies. Sensors 20 (9), 2700.

Just, W., 2001. Computational complexity of multiple sequence alignment with SP-score. J. Comput. Biol. 8 (6), 615–623.

Keogh, E.J., Pazzani, M.J., 2001. Derivative dynamic time warping. In: Proceedings of the 2001 SIAM International Conference on Data Mining. SIAM, pp. 1–11.

Kotas, M., Pander, T., Leski, J.M., 2015. Averaging of nonlinearly aligned signal cycles for noise suppression. Biomed. Signal Process. Control 21, 157–168.

Li, H., Liu, J., Yang, Z., Liu, R.W., Wu, K., Wan, Y., 2020. Adaptively constrained dynamic time warping for time series classification and clustering. Inform. Sci. 534, 97–116.

Liu, Y.-T., Zhang, Y.-A., Zeng, M., 2018. Novel algorithm for hand gesture recognition utilizing a wrist-worn inertial sensor. IEEE Sens. J. 18 (24), 10085–10095.

Liu, Y.-T., Zhang, Y.-A., Zeng, M., 2019. Adaptive global time sequence averaging method using dynamic time warping. IEEE Trans. Signal Process. 67 (8), 2129–2142.

Łuczak, M., 2016. Hierarchical clustering of time series data with parametric derivative dynamic time warping. Expert Syst. Appl. 62, 116–130.

MathWorks, 2019. Modified Akima piecewise cubic Hermite interpolation. https://ww2.mathworks.cn/help/matlab/ref/makima.html?lang=en.

Maus, V., Câmara, G., Cartaxo, R., Sanchez, A., Ramos, F.M., De Queiroz, G.R., 2016. A time-weighted dynamic time warping method for land-use and land-cover mapping. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (8), 3729–3739.

Meesrikamolkul, W., Niennattrakul, V., Ratanamahatana, C.A., 2012. Shape-based clustering for time series data. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, pp. 530–541.

Morel, M., Achard, C., Kulpa, R., Dubuisson, S., 2018. Time-series averaging using constrained dynamic time warping with tolerance. Pattern Recognit. 74, 77–89.

Mudelsee, M., 2019. Trend analysis of climate time series: A review of methods. Earth-Sci. Rev. 190, 310–322.

Needleman, S.B., Wunsch, C.D., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48 (3), 443–453.

Niennattrakul, V., Ratanamahatana, C.A., 2009. Shape averaging under time warping. In: 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Vol. 2. IEEE, pp. 626–629.

Niennattrakul, V., Srisai, D., Ratanamahatana, C.A., 2012. Shape-based template matching for time series data. Knowl.-Based Syst. 26, 1–8.

Okawa, M., 2019. Template matching using time-series averaging and DTW with dependent warping for online signature verification. IEEE Access 7, 81010–81019.

Okawa, M., 2020. Online signature verification using single-template matching with time-series averaging and gradient boosting. Pattern Recognit. 102, 107227.

Okawa, M., 2021. Time-series averaging and local stability-weighted dynamic time warping for online signature verification. Pattern Recognit. 112, 107699.

Petitjean, F., Forestier, G., Webb, G.I., Nicholson, A.E., Chen, Y., Keogh, E., 2014. Dynamic time warping averaging of time series allows faster and more accurate classification. In: 2014 IEEE International Conference on Data Mining. IEEE, pp. 470–479.

Petitjean, F., Forestier, G., Webb, G.I., Nicholson, A.E., Chen, Y., Keogh, E., 2016. Faster and more accurate classification of time series by exploiting a novel dynamic time warping averaging algorithm. Knowl. Inf. Syst. 47 (1), 1–26.

Petitjean, F., Gançarski, P., 2012. Summarizing a set of time series by averaging: From steiner sequence to compact multiple alignment. Theoret. Comput. Sci. 414 (1), 76–91.

Petitjean, F., Ketterlin, A., Gançarski, P., 2011. A global averaging method for dynamic time warping, with applications to clustering. Pattern Recognit. 44 (3), 678–693.

Redmond, D.P., Sing, H.C., Hegge, F.W., 2020. Biological time series analysis using complex demodulation. In: Rhythmic Aspects of Behavior. Routledge, pp. 429–457.

Sakoe, H., Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. Speech Signal Process. 26 (1), 43–49.

Salgotra, R., Gandomi, M., Gandomi, A.H., 2020. Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming. Chaos Solitons Fractals 138, 109945.

Schultz, D., Jain, B., 2018. Nonsmooth analysis and subgradient methods for averaging in dynamic time warping spaces. Pattern Recognit. 74, 340–358.

Sioros, G., Nymoen, K., 2021. Accurate shape and phase averaging of time series through dynamic time warping. arXiv preprint arXiv:2109.00978.

Sun, T., Liu, H., Yu, H., Chen, C.P., 2016. Degree-pruning dynamic programming approaches to central time series minimizing dynamic time warping distance. IEEE Trans. Cybern. 47 (7), 1719–1729.

Tran, T.M., Le, X.-M.T., Nguyen, H.T., Huynh, V.-N., 2019. A novel non-parametric method for time series classification based on k-nearest neighbors and dynamic time warping barycenter averaging. Eng. Appl. Artif. Intell. 78, 173–185.

Wang, L., Jiang, T., 1994. On the complexity of multiple sequence alignment. J. Comput. Biol. 1 (4), 337–348.

Zhao, J., Itti, L., 2018. shapeDTW: Shape dynamic time warping. Pattern Recognit. 74, 171–184.

Zhu, L., Najafizadeh, L., 2017. Dynamic time warping-based averaging framework for functional near-infrared spectroscopy brain imaging studies. J. Biomed. Opt. 22 (6), 066011.