Pattern Recognition 81 (2018) 268–279

Contents lists available at ScienceDirect

Pattern Recognition

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

Time Alignment Measurement for Time Series

Duarte Folgado[a,*], Marília Barandas[a], Ricardo Matias[b,c], Rodrigo Martins[b], Miguel Carvalho[d], Hugo Gamboa[e]

[a] Associação Fraunhofer Portugal Research, Rua Alfredo Allen 455/461, Porto, Portugal
[b] Physiotherapy Department, School of Health, Polytechnic Institute of Setúbal, Estefanilha, Edifício da ESCE, 2914-503, Setúbal, Portugal
[c] Champalimaud Research. Champalimaud Centre for the Unknown, Lisbon, Portugal
[d] Minho University, Campus de Azurém, 4800-058, Guimarães, Portugal
[e] Laboratório de Instrumentação, Engenharia Biomédica e Física da Radiação (LIBPhys-UNL), Departamento de Física, Faculdade de Ciências eTecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Article history:
Received 26 July 2017
Revised 3 January 2018
Accepted 2 April 2018
Available online 3 April 2018

Keywords:
Time series
Time warping
Similarity
Distance
Signal alignment

ABSTRACT

When a comparison between time series is required, measurement functions provide meaningful scores to characterize similarity between sequences. Quite often, time series appear warped in time, i.e, although they may exhibit amplitude and shape similarity, they appear dephased in time. The most common algorithm to overcome this challenge is the Dynamic Time Warping, which aligns each sequence prior establishing distance measurements. However, Dynamic Time Warping takes only into account amplitude similarity. A distance which characterizes the degree of time warping between two sequences can deliver new insights for applications where the timing factor is essential, such well-defined movements during sports or rehabilitation exercises. We propose a novel measurement called Time Alignment Measurement, which delivers similarity information on the temporal domain. We demonstrate the potential of our approach in measuring performance of time series alignment methodologies and in the characterization of synthetic and real time series data acquired during human movement.

1. Introduction

The comparison of time series have existed in the scenario of sequence matching, subsequence searching, and motif detection. Those challenges are intrinsically related to time series classification applied in several contexts such as pattern recognition [1–4], signal processing [5], shape detection [6], bioinformatics [7,8], human activity recognition [9] and on-line handwritten signature validation [10].

When a comparison of two streams of data with implicit or explicit time information associated is executed, there is the need for a measurement function that provides information on the similarity of the two data streams. Time series comparison may be established using a wide range of available distance measurement functions. Some of the traditional metrics, such the Euclidean distance or some modification thereof, assume that the discrete signals are equidistant points in time and also aligned in the time axis. In some domains, although time series may present amplitude and shape similarity, they can be considered to be out-of-phase. Therefore, similar regions may appear in different instants in time, leading to different degrees of time distortion, or time warping, among several sequences, since they are not aligned in the temporal domain. In those circumstances, traditional distances fail to measure this distortion since they are very sensitive to small distortions in time and typically unable to directly handle unequal length time series without some sort of preprocessing [11].

In order to overcome these limitations, elastic distances which contemplate temporal elastic shifting have been proposed. Dynamic Time Warping (DTW) and Longest Common Subsequence (LCSS) compensate non-linear temporal distortions by aligning the discrete sequences before establishing amplitude measurements in the discrete domain [12]. Since those algorithms do not take into account the information between inter-sampling points, [13] proposed the Continuous Dynamic Time Warping (CDTW), which extends the classic methodology by allowing mapping between instants that may eventually not belong to the original time vector for each series. The work from [14] uses an optimization approach to calculate a parametric polynomial warping path reflecting the alignment between both series. Therefore, the last two alternatives

* Corresponding author.
E-mail address: duarte.folgado@fraunhofer.pt (D. Folgado).

https://doi.org/10.1016/j.patcog.2018.04.003

produce an optimal warping path which translates the alignment between two signals in the continuous domain.

Motivated by the fact that off-the-shelf applications of semi-supervised learning algorithms do not typically work well when applied to time series, the authors from [15] proposed a new distance which tries to minimize this behaviour. The proposed distance is called Dynamic Time Warping Delta (DTW-D) and is the ratio between DTW and Euclidean distances.

Inspired by the well-known *edit distance* for string comparison, which calculates the minimum number of insertions, deletions, and substitution operations to transform a string in another, some authors translated the core idea to time series [16–18]. In order to generalize the concept from strings to time series, two elements of each sequence are matched if the absolute difference between them is bellow a given tolerance value. The common goal of the approaches consists in identifying the smallest number of operations (additions, deletions and substitutions) to transform a sequence in another.

Prior to establish a similarity measurement between time series, most of the aforementioned examples perform a previous alignment between the two sequences. The optimal alignment may also be used for summarizing a set of time series, since it allows to compute a more meaningful average between sequences which may exhibit time warping. The work developed by [19], and more recently by [20], proposes time series averaging methods based on preceding alignments, which demonstrated favourable impacts on clustering performance.

However, whilst we observed a multitude of proposed novel elastic distances over the last years, they are mostly centered in measuring similarity accounting for amplitude differences [21,22]. Those facts motivated our work in the development of a novel time distance able to measure similarity between time series in the temporal domain, namely Time Alignment Measurement (TAM). The proposed methodology is able to describe the behaviour in time between two signals by measuring the fraction of time distortion between them. The distortion may comprise periods of temporal advance or periods of delay. When signals are similar-alike in time they can be considered to be in phase between each other. This approach can deliver useful information to domains where information between the temporal misalignment of time series is needed. Examples of such domains include well-defined human movements executed in sports or rehabilitation exercises. The authors from [23] investigated the feasibility of biofeedback training applied to therapeutic exercises, where repetitive movements should follow well-defined timings to be considered successfully executed. The authors calculated the mean error of the distance between anatomic segments executed by the subject to a previously recorded reference. A distance able to truly characterize temporal misalignment between movements should bring new perspectives for the evaluation of the correctness of the exercises through the complete movement execution.

The literature review allowed to identify that most of the work developed over the last years in the development of new distance functions mostly takes into account amplitude similarity. The major contribution presented on this work is propose a novel distance which measures similarity in time domain.

The remaining content of this paper is organized as follows: in section 2, a brief overview of DTW algorithm is presented, since we use DTW to align two time series prior calculating TAM. Section 3 introduces the TAM distance and presents examples based on synthetic time series to support its potential. In section 4 we present two use cases for the proposed distance based on real time series data. Finally, section 5 contains the conclusions and future work directions.

2. Time series alignment

In this section, we motivate for the utility of DTW algorithm to establish an alignment between two time series in order to calculate TAM. We start with a brief explanation of DTW algorithm and explore some of the challenges arising while aligning signals that present amplitude fluctuation.

2.1. Dynamic Time Warping

The DTW algorithm allows two time-dependent sequences that are similar, but locally out of phase, to align in time. Its main objective consist of identifying an optimal alignment between sequences by warping the time axis iteratively.

In order to align two time series $X:=(x_1, x_2, ..., x_N)$ and $Y:=(y_1, y_2, ..., y_M)$ of length N and M respectively, a N-by-M *cost matrix* is computed. Each (n^{th}, m^{th}) element of the *cost matrix*, $C \in \mathbb{R}^{N\times M}$, corresponds to the distance between each pair of elements of the sequences X and Y. The Euclidean distance is usually employed as a distance function to define the *cost matrix* element as:

$$c(x_n, y_m) = (x_n - y_m)^2 \tag{1}$$

The goal of DTW is to find the optimal warping alignment path between X and Y having minimum overall cost. A warping path, W, is a set of matrix elements that define the relationship between X and Y. The k^{th} element of W is defined as $w_k = (i, j)_k$, $w_k \in \mathbb{R}^2$:

$$W = (w_1, w_2, ..., w_k, ..., w_K) \qquad \max(N, M) \le K \le N + M - 1 \tag{2}$$

The resulted path should be composed by a set of matrix elements satisfying the following conditions:

- **Boundary condition**: Enforces that the first and the last elements of X and Y are aligned to each other $\therefore w_1 = (1, 1)$ and $w_K = (N, M)$.
- **Monotonicity condition**: Forces the points in the warping path to be monotonically spaced in time $\therefore i_1 \le i_2 \le ... \le i_N$ and $j_1 \le j_2 \le ... \le j_M$.
- **Step size condition**: Avoids omissions in elements and replications in the alignment of X and Y $\therefore (w_{k+1} - w_k) \in (1, 0), (0, 1), (1, 1)$ for $k \in [1 : K - 1]$.

The optimal warping path is the path that has the minimum total cost among all possible warping paths. One could test every incumbent warping path and determine the minimum cost candidate, but such method will lead to a exponential computational complexity in the lengths of N and M. Using dynamic programming, an accumulated cost matrix, D, is computed in order to find the path that minimizes the warping cost in an $O(N, M)$ complexity [12]. Each accumulated cost matrix element is defined as the local cost measure in the current cell plus the minimum of the local cost measures in the adjacent cells:

$$D(n, m) = min\{D(n-1, m-1), D(n-1, m), D(n, m-1)\} + c(x_n, y_m) \tag{3}$$

where $n \in [1: N]$, $m \in [1: M]$, D is the accumulated cost matrix, and $c(x_n, y_n)$ is the local cost measure found in the current cell.

Using this accumulated matrix, the optimal warping path, $W^* = (w_1, w_2, ..., w_K)$, is computed in reverse order of indices, starting with $w_K = (N, M)$, by the following algorithm:

$$w_{k-1} = \begin{cases} (1, m-1), & \text{if } n = 1 \\ (n-1, 1), & \text{if } m = 1 \\ argmin\{D(n-1, m-1), D(n-1, m), D(n, m-1)\}, & \text{otherwise} \end{cases} \tag{4}$$

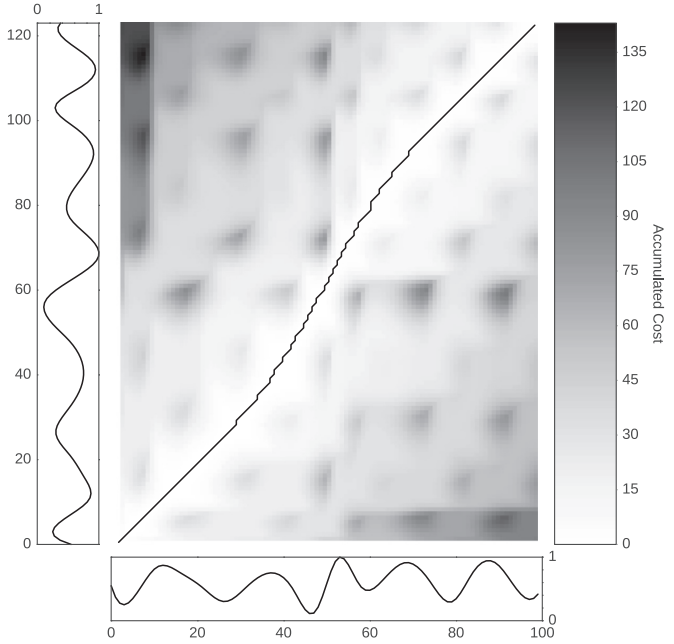

Fig. 1. Accumulated cost matrix between two time series using the Euclidean distance as local cost measure. The resulted optimal warped follows the low cost regions (represented in white) and avoids high cost regions (represented in dark).

In order to compensate the effect of different optimal warping path lengths, the path-normalized distance is given as:

$$DTW(X,Y) = \frac{1}{K}\sqrt{\sum_{k=1}^{K} w_k^*} \tag{5}$$

Fig. 1 illustrates a typical example of DTW algorithm. The reference time vector of the lower bottom signal was artificially modified to result in the time warped signal represented vertically. The resulting optimal path follows the minimum cost regions on the accumulated cost matrix and establishes a pairwise relationship between each point of the discrete series.

DTW distance yields to a more intuitive information on signal amplitude by performing a preceding alignment before calculating the distance between two signals. Additionally, the resulted optimal alignment path settles a pairwise relationship between each point of the discrete signals. This discrete temporal matching may be used for signal alignment. In our work, we further explored this feature by creating a distance to measure time distortions between two signals.

2.2. Signal alignment challenges

Although DTW has been successfully used for many years, it still encounters some pairwise alignment challenges. In [24] the authors reported unintuitive alignments when the algorithm tries to express amplitude variability in the Y-axis by improper warping the X-axis. This behaviour leads to situations defined as "singularities", where a single point of a particular signal maps a large subsection of another time series. In order to overcome the singularities challenge they presented the Derivative Dynamic Time Warping (DDTW) approach, which uses the square of the differences between the estimated signal derivatives as shown on Fig. 2.

Despite the fact this methodology reduces the number of singularities and does not completely solve the problem, it has been successfully used in many fields, including human activity recognition using accelerometer signals [26] and biosignal segmentation [27]. However, since DDTW uses the first signal derivative it is also quite susceptible to noise and requires signal smoothing before applying the algorithm.

In [28] the authors introduced a DTW penalty based version called Weighted Dynamic Time Warping (WDTW). In this approach, the *cost matrix* is modified in order to incorporate a modified logistic weight function that assigns additional weight as a function of the phase difference between the reference and test points. Thus, time instants with higher phase difference will be more penalized than instants near the reference.

More recently, [29] presented an approach that solely accounts for the shape of the time series. The similarity measure is performed by comparing the spatial distribution of the data around each point. This modification tends to reduce singularities and promotes feature alignment that may include peaks and valleys.

Despite the attempts to improve the DTW alignment they are still dependent of the data's nature. For instance, time series that do not comprise higher degree of information on the first derivative are susceptible not to benefit from an alignment solely based on the data's shape. Therefore, an alignment which contemplates a weighting between the amplitude and derivative domains can constitute an added value towards a more versatile application.

2.3. Sliding Window Dynamic Time Warping

In order to overcome the incorrect alignment generated by DTW and its variants, we outlined an alternative alignment that should prevent singularities by reflecting a feature-to-feature similarity. We defined features as notable events on the course of time series, which include local minima, maxima or valleys. Therefore, features should always be aligned with the corresponding features in the other signal. Furthermore, a feature should always correspond to a single point, since by definition they are unique in time. The proposed approach is called Sliding Window Dynamic Time Warping (SW-DTW).

The cost measured used by DTW uses an element-to-element distance. In order to use a "contextual" distance, which takes into account the neighbourhood surrounding each point, we modified the cost function to take into account the distance between well-defined windows of the signals. The modified local cost measure can be defined as:

$$\begin{aligned} c(x_n, y_m) &= w(\delta) \times \mathbf{W}(\delta) \\ \mathbf{W}(\delta) &= \alpha(X_{[n-\frac{\delta}{2}:n+\frac{\delta}{2}]} - Y_{[n-\frac{\delta}{2}:n+\frac{\delta}{2}]})^2 \\ &\quad + (1-\alpha)(\Delta[X]_{[m-\frac{\delta}{2}:m+\frac{\delta}{2}]} - \Delta[Y]_{[m-\frac{\delta}{2}:m+\frac{\delta}{2}]})^2 \end{aligned} \tag{6}$$

where $w(\delta)$ is a window function with width $\delta \in \mathbb{N}$, Δ is the operator for the first discrete derivative, $\alpha \in \mathbb{R}_0^+ \cap [0:1]$ is a constant that defines the weighting between the cost in amplitude and first order derivative, n and m are the indexes of the values of X and Y, respectively. Using a Hanning window function we can assure that points closer to the center of the window have more contribution to the local cost than points located near the window limits. Since the first discrete derivative of the signal is calculated, the last elements of X and Y are discarded to guarantee that both time series have the same length. Additionally, signals are prepared by introducing reflected copies of the each signal (with the window size) in both ends. This procedures aims to minimize boundary errors on the first and last elements of each signal. δ and α are free parameters and, consequently, must be tuned prior to applying the algorithm. Small windows will tend to similar results compared with the point-to-point DTW distance and excessive large window size will tend to improper feature alignment.

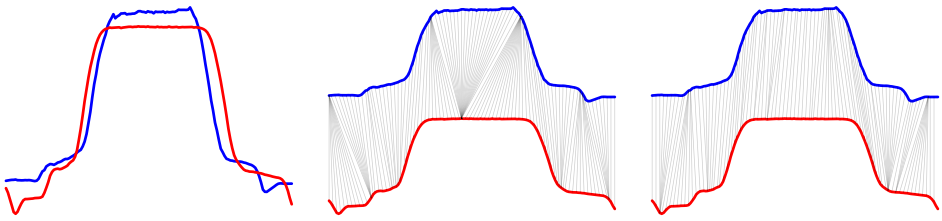

Fig. 2. Two time series from Gun Point dataset [25] aligned with the DTW approach (center) and DDTW modification (right). Although a slight improvement can be observed, there are still sections of consecutive singularities.

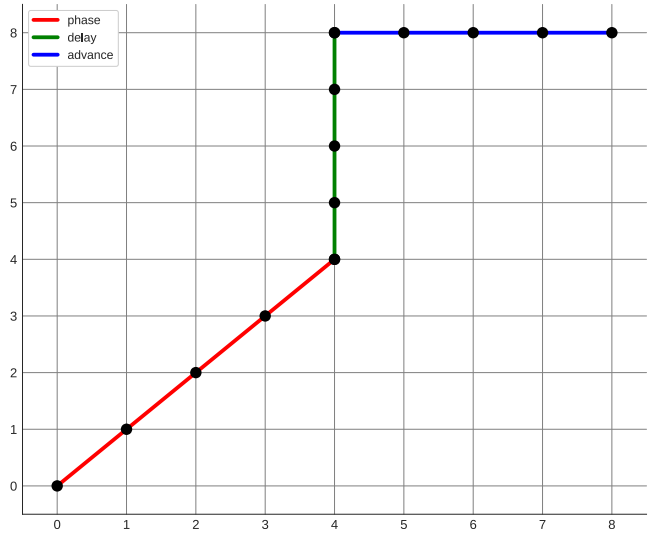

Fig. 3. An example of an artificial optimal warping path superimposed on an accumulated cost matrix containing all the available step directions. The horizontal, vertical and diagonal segments represent the advance, delay and phase intervals, respectively.

3. Time Alignment Measurement

In this section we will start to present a further interpretation of the optimal warping path from the DTW algorithm. This analysis will allow to easily present the TAM formulation. We will finish this section with illustrative examples to simultaneously consolidate the presentation and support the potential of our proposed distance.

3.1. DTW optimal alignment path properties

The optimal warping path establishes a pairwise relationship between the indexes of both series. This resemblance allows to characterize the transformations on the time axis between them. Whilst there are a multitude of step patterns proposed on the literature for the warping path calculation, we started by exploring the basic step pattern which contemplates vertical, horizontal and diagonal segments that was discussed in section 2.1.

Let us consider that the eventual time warping is referenced to series X which is plotted on the x-axis. The possible slopes which are contemplated on the warping are outlined on Fig. 3.

Horizontal segments.

An horizontal segment is defined as $w_{k+1} - w_k = (1, 0)$. In this case, an index of Y is associated to one or more consecutive indexes of the reference X. This situation illustrates a temporal advance as the same time instant is maintained on the Y sequence, several time instants are elapsing in the reference series.

Vertical segments.

A vertical segment can be defined as $w_{k+1} - w_k = (0, 1)$. This situation arises when an index of X is associated to one or more consecutive indexes of the series Y. A temporal delay is therefore present since sequence Y is progressing in time and the reference maintains the same instant.

Diagonal segments.

A diagonal segment is defined as $w_{k+1} - w_k = (1, 1)$. In this circumstance, there is no time warping and the signals can be considered to display phase phenomenon between them.

3.2. Outline

The idea behind the proposed distance is to measure the cost between a given time series to warp in time relative to the other. Using the optimal alignment path between two time series we can extract information in the time domain that allows to characterize the intervals when the series are in phase, advance or delay.

Let us consider again two sequences X of length $N \in \mathbb{N}$ and Y of length $M \in \mathbb{N}$. During the complete length of each sequence, the signals may be considered to exhibit phase and out of phase behaviours. In case the signals are out of phase, one sequence can be considered to be in advance in relation to the other. This characteristic is reciprocal as if X is in advance in relation to Y, Y can be considered to be delayed in relation to X.

If we assume that Y is delayed in relation to X, the total time which Y is delayed in relation to X is denoted as $\overleftarrow{\theta}_{xy}$ and the time which may be eventually advanced is denoted as $\overrightarrow{\theta}_{xy}$. Using this nomenclature, we can write the relation between advance, delay, and length of both signals as:

$$|\overrightarrow{\theta_{xy}} - \overleftarrow{\theta_{xy}}| = |M - N| \tag{7}$$

The total time when both signals are in phase is represented by $\overline{\theta_{xy}}$. During the complete length of signal Y, the fraction of advance ($\overrightarrow{\psi}$), delay ($\overleftarrow{\psi}$), and phase ($\overline{\psi}$) to X can be calculated as:

$$\overrightarrow{\psi} = \frac{\overrightarrow{\theta_{xy}}}{N} \qquad \overleftarrow{\psi} = \frac{\overleftarrow{\theta_{xy}}}{M} \qquad \overline{\psi} = \frac{\overline{\theta_{xy}}}{\min\{N, M\}} \tag{8}$$

Finally, the TAM distance can be formally defined as:

$$\Gamma = \overrightarrow{\psi} + \overleftarrow{\psi} + (1 - \overline{\psi}), \qquad \Gamma \in \{\mathbb{R}_0^+ | \Gamma \in [0:3]\} \tag{9}$$

This distance penalizes signals where advance or delay is present and benefits series that are in phase between each other. As the distance increases, the dissimilarity between both signals also increases. Thus, in case the signals are constantly in phase, $\overline{\psi} = 1$, $\overrightarrow{\theta_{xy}} = 0$, $\overrightarrow{\psi} = 0$, $\overleftarrow{\theta_{xy}} = 0$ and $\overleftarrow{\psi} = 0$. The TAM distance has the minimum allowed value ($\Gamma = 0$), and the signals can be considered equal in temporal domain. The highest dissimilarity value is traduced when $\Gamma = 3$, where $\overrightarrow{\theta_{xy}} = N \Rightarrow \overrightarrow{\psi} = 1$, $\overleftarrow{\theta_{xy}} = M \Rightarrow \overleftarrow{\psi} = 1$, and consequently $\overline{\psi} = 0$.

It is important to note some considerations regarding the topology of the proposed measure:

1. The condition of identity of indiscernibles is not satisfied: $\Gamma(X,Y)=0 \nRightarrow X=Y$. In fact, two signals can be equal in time and possess dissimilarity in amplitude. A trivial application of our distance to two similar signals in time with an offset on amplitude proves this assumption.
2. Symmetry is observed since the concept of advance and delay between two time series is reciprocal: $\overrightarrow{\theta}_{xy}=\overleftarrow{\theta}_{yx}=\alpha$ and $\overleftarrow{\theta}_{xy}=\overrightarrow{\theta}_{yx}=\beta$. Furthermore, $\overline{\theta_{xy}}=\overline{\theta_{yx}}=\overline{\theta}$. The symmetry proof is trivial and outlined on Eq. 10:

$$\begin{aligned}&\frac{\overrightarrow{\theta}_{xy}}{N}+\frac{\overleftarrow{\theta}_{xy}}{M}+\left(1-\frac{\overline{\theta_{x,y}}}{\min(N,M)}\right)\\&\quad=\frac{\overrightarrow{\theta}_{yx}}{N}+\frac{\overleftarrow{\theta}_{yx}}{M}+\left(1-\frac{\overline{\theta_{yx}}}{\min(N,M)}\right)\Leftrightarrow\\&\quad\Leftrightarrow\frac{\alpha}{N}+\frac{\beta}{M}+\left(1-\frac{\overline{\theta}}{\min(N,M)}\right)\\&\quad=\frac{\beta}{M}+\frac{\alpha}{N}+\left(1-\frac{\overline{\theta}}{\min(N,M)}\right)\Leftrightarrow\\&\quad\Leftrightarrow\Gamma(X,Y)=\Gamma(Y,X)\end{aligned}\tag{10}$$

The aforementioned considerations reveal that TAM can not be considered as a metric since it fails to guarantee the identity of indiscernibles. However, we can state that is a *premetric*, since it fully satisfies both the non-negativity and symmetry conditions. It is important to empathize that in order to calculate the TAM distance, it is only required to establish a pairwise relation between the elements of each time series. This pairwise relation provides the required alignment to compute the delays, advances, and phase periods between signals. Thus, alternative methods for signal alignment to DTW can also be used to compute TAM.

Calculate TAM from optimal warping path.

TAM can be calculated directly from the DTW warping path based on the following assumptions:

Let $\Delta[W_k^*]=w_{k+1}^*-w_k^*$ be the finite difference between two consecutive coordinates of the optimal warping path at point k represented as a bidimensional vector. Since the optimal warping path is restricted to vertical, horizontal and diagonal segments, $\Delta[W_k^*]$ is also rescricted to the values of $\Delta[W_k^*]\in\{(1,1),(1,0),(0,1)\}$. The vertical segments will have a value of $\Delta[W_k^*]=(0,1)$, marked by a temporal delay; the horizontal segments will have $\Delta[W_k^*]=(1,0)$, denoting a temporal advance, and the diagonal segments will present $\Delta[W_k^*]=(1,1)$, denoting phase among the time series.

The number of instants in advance, delay, and phase can be directly calculated from the optimal path accordind to Eq. 11.

$$\begin{aligned}\overrightarrow{\delta_k}&=\begin{cases}1, & \Delta[W_k^*]=(1,0)\\0, & \text{otherwise}\end{cases}\\\overleftarrow{\delta_k}&=\begin{cases}1, & \Delta[W_k^*]=(0,1)\\0, & \text{otherwise}\end{cases}\\\overline{\delta_k}&=\begin{cases}1, & \Delta[W_k^*]=(1,1)\\0, & \text{otherwise}\end{cases}\end{aligned}\tag{11}$$

Hence, TAM can be calculated according to Eq. 12:

$$\Gamma=\frac{1}{N}\sum_{k=1}^{K}\overrightarrow{\delta_k}+\frac{1}{M}\sum_{k=1}^{K}\overleftarrow{\delta_k}+\left(1-\frac{1}{\min\{N,M\}}\sum_{k=1}^{K}\overline{\delta_k}\right)\tag{12}$$

Returning to the analysis of Fig. 3 we can see that the path starts within phase, enters in delay during four segments and finishes in advance during another four segments. We can write the resultant $\overrightarrow{\delta}$, $\overleftarrow{\delta}$, and $\overline{\delta}$ as:

$$\begin{aligned}\overrightarrow{\delta}&=\{0,0,0,0,0,0,0,0,1,1,1,1\}\\\overleftarrow{\delta}&=\{0,0,0,0,1,1,1,1,0,0,0,0\}\\\overline{\delta}&=\{1,1,1,1,0,0,0,0,0,0,0,0\}\end{aligned}\tag{13}$$

3.3. Application

In the following paragraphs we will use artificial signals to explain the proposed distance and support its potential to characterize time series.

Equal length time series.

When we establish a comparison between two time series, X and Y, with equal lengths $N=M=L$, the duration of the intervals where X is in advance and delay in comparison to Y must be equal. This property is a consequence of the fact that both signals have equal lengths. Therefore, despite an eventual delay, the signal must have at least an *a posteriori* advance to finish at the same instant as the other sequence and vice-versa.

Since $\overrightarrow{\theta_{xy}}=\overleftarrow{\theta_{xy}}=\theta$, the TAM distance can be directly simplified to:

$$\Gamma(X,Y)=\frac{3\theta}{L}\tag{14}$$

Fig. 4 represents an example where a set of four artificial signals was generated by distorting the first sequence. Although in this example the Euclidean distance increases with the increase of time distortion, it does not reflect a meaningful measurement. Since it is only sensitive to amplitude similarity, in case a sequence possess an offset on the plateau values, it will produce a greater distance value, independently of the time warping degree. The DTW aligns each pair of signals prior computing the distance and thus produces equal scores for all examples. However, TAM produces a meaningful score, which reflects the cost to compress and dilate in time a specific signal. In fact, the distance increases with the degree of time distortion present in each sequence. Signal **A** is identical to the reference signal. The plateau from signal **B** has an advance of 10 seconds. In order to finish at the same instant, the signal enters in delay at the last segment for another 10 seconds. Signal **C** is distorted half of its time as $\overrightarrow{\theta_{xy}}=\overleftarrow{\theta_{xy}}=15\Rightarrow\Gamma=\frac{3\cdot15}{30}$. Signal **D** represents a significant advance of the plateau, which is represented as a single peak, followed by a significant delay until the end of the signal. Note that in all examples the two first and last samples are in phase, which prevents to achieve the maximum theoretical TAM value.

Unequal length time series.

Time series with different lengths may present multiple behaviours in the way they perform in time. Fig. 5 represents a group of four signals distorted in time. The signals were generated from the first sequence by modifying the respective time vectors to simulate delays and advances. Signal **A** shares the same time representation with the reference signal. Signal **B** possess an initial delay of 10 seconds then enters in phase for the rest of its length. The TAM distance of 0.25 reflects this compression. A higher distance is observed for signal **C** since a more significant advance is present. Signal **D** reflects a linear delay lag since two instants in the signal are related to a single instant in signal **A**, producing a distance value of 0.5. The Euclidean distance was calculated by linear interpolating all signals to share the same length of the reference. This procedure assures that signals have equal lengths in order to compute the Euclidean distance. Although signal **D** is different from the reference in time domain, the Euclidean distance is zero.

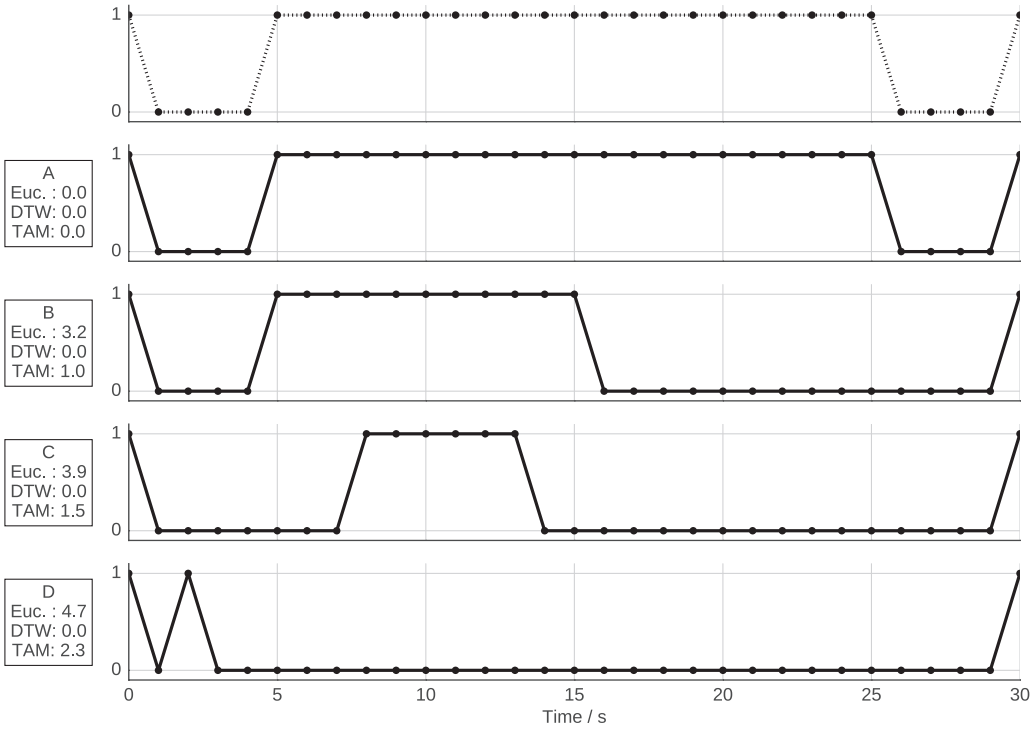

Fig. 4. A set of four examples with equal lengths. The first sequence is the reference signal (dashed line). All the signals are compared against the reference signal. An annotation is provided with each respective Euclidean, DTW and TAM distances.

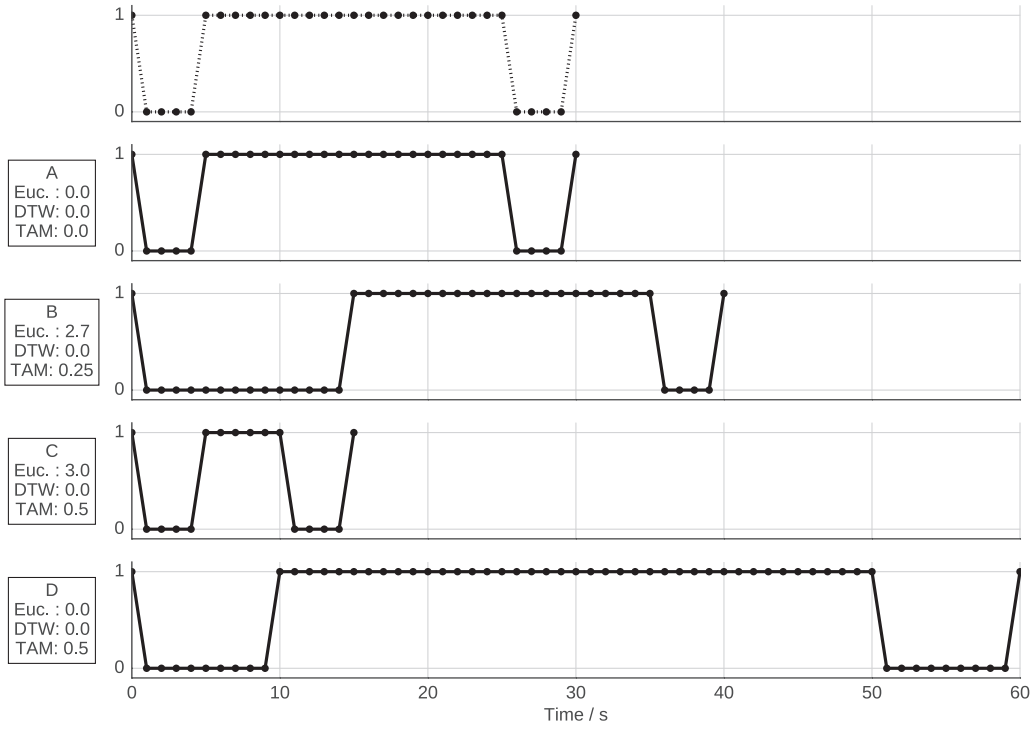

Fig. 5. A group of four signals distorted in time. The signals were generated from the upper signal (dashed line) by artificially modifying the vectors to simulate delays and advances. An annotation is provided to show the Euclidean, DTW and TAM distances. The Euclidean distance was calculated by linear interpolating all signals to the reference length.

Application notes.

The previous examples allowed to better describe the nature of our main contribution: provide a novel distance measurement able to characterize the degree of time warping between time series which may be similar-alike in amplitude. The DTW-D, proposed by [15], consists of the ratio between DTW and the Euclidean distance. One might potentially argue that such ratio could be an approximation to measure warping, as it measures the amount of warping necessary to match a given time series in reference to the Euclidean distance (which requires no warping at all). However,

DTW-D will eventually fail in the presented application examples. As the signals are similar-alike in amplitude, DTW will have the value of 0 and, consequently, DTW-D will also fail to provide a meaningful score.

As a final note, a naive approach to compare time series based on the time domain would be solely compare the length of the sequences. However, the TAM evaluates the temporal behaviour in terms of delay, advance and phase along the time of each sequence and, therefore, it not strictly limited to the endpoints of each sequence. Consequently, even for sequences with equal length, temporal information can be extracted which would not be possible if a direct comparison between signal lengths was performed.

4. Experimental Evaluation

In this section two studies will be presented to demonstrate the applicability and relevance of our approach to characterize real time series data. As previously mentioned in subsection 3.2, the TAM value is calculated based on the previous time series alignment. Therefore, the value depends on the preceding alignment quality. The first study consists of examining the signal alignment quality using well-known DTW variations and our proposed SW-DTW modification. Secondly, we will apply the TAM as a local measure to examine human repetitive motion using inertial data.

4.1. Simulated time series alignment

We created a controlled experiment in order to assess the signal alignment performance across several DTW variations. During the course of our research, we did not find a dataset whose main objective is to serve as validation for time series alignment mechanisms. In this sense, we implemented a study based upon a comparison between a given time series X and a modified time series $\hat{X}$ calculated from an amplitude modification of X.

A scale vector, S, was generated using a series of random values from a Gaussian distribution. In order to prevent an excessive modification between consecutive elements, we used a similar approach to [30], where the initial random values were filtered to ensure adjacent scales differ by at most 1: $S(t+1) = S(t) + \sin(\pi \times randn)$. The signal was multiplied by the scale vector in order to modulate negative and positive fluctuations: $\hat{X} = X \otimes S$. This procedure results in two time series which are always in phase during their entire length, since the unique modification was implemented in the amplitude domain (taking also into account that no excessive modification was performed in order to prevent significant changes in the shape of the two signals).

When using DTW and its variants, the ideal expected outcome is an optimal warping path which demonstrates that the signals are continuously in phase during their complete length. However, the amplitude fluctuations arising from the multiplication with the scale vector are susceptible to generate singularities as previously discussed in subsection 2.2.

In order to quantify the alignment quality, the TAM was calculated between each pair of time series. Since the signals are aligned throughout their complete length, it is expected that $\Gamma = 0$ in circumstances where the alignment was indeed performed correctly. Given that singularities result in advances and delays that do not correspond to correct alignments, the value of the TAM will be incorrectly influenced. Therefore, in the context of this experiment, the TAM value can be used to translate the alignment quality and establish a comparison among different signal alignment techniques.

We used the UCR time series archive [25] to test several DTW variations by randomly selecting 40 signals from 84 datasets, which resulted in 3360 different alignments per algorithm variation. The selected algorithms were the DTW (no warping window), the DTW with a 5% warping window of signal's length (DTW_R), the DDTW, and the SW-DTW with empirical values of $\alpha = 0.5$ and $\delta = 0.05 \times N$. The scale vector, S, was normalized prior the multiplication in order to guarantee that $S \in [0.5, 1.5]$. The results are summarized in Table 1.

The analysis suggests that SW-DTW reduces the number of singularities as it outperforms the other variants in most of the datasets. The improvement of DTW_R in comparison with DTW is explained by the fact the maximum distance of the warping path to the diagonal is restricted. In the majority of the situations where the SW-DTW is not the best alignment alternative for a given dataset, the lowest Γ is achieved by the DDTW. It is important to emphasize that the value of $\alpha = 0.5$ was used for all datasets and that no individual adjustment was performed in order to reduce the complexity of the analysis. Since lower values of α will increase the weight of the first order derivative, we can anticipate that it can be used to increase the alignment quality by SW-DTW in datasets where the DDTW achieved superior performance. This fact also suggests that before applying SW-DTW, proper tuning of the α and δ parameters is required.

The different DTW alignment methodologies resulted in different alignments for the same dataset and express variability in the Γ values. Therefore, we can anticipate that these results support our claim that TAM is sensitive to the alignment quality and that SW-DTW reduces the singularities in comparison with the other evaluated alternatives. It is worth to mention that despite SW-DTW achieved superior performance on this experiment, it is not our main contribution. Since TAM depends on preceding alignments, supported by the results of Table 1, this experiment allowed to increase our confidence that SW-DTW reduces the number of singularities and produces a more correct alignment in comparison with the evaluated alternatives.

A detailed analysis of the UCR dataset also allowed to elaborate important highlights before attempting to proceed with a time series classification exercise using TAM as a local measurement. The TAM should be used in datasets with significant temporal distortion and similar amplitude between different classes. Additionally, each class must also comprise time series which are similar-alike in the temporal domain. This situation is not present in the majority of the UCR datasets since there are several datasets with minor temporal differences between classes (e.g. *Adiac, OliveOil* and *ProximalPhalanxOutlineCorrect*). Therefore, we introduced a new time series dataset that suits the TAM applicability requirements and will be thoroughly discussed in subsection 4.2.

4.2. Repetitive upper limb motion

A base motivation for the development of this new measure was to describe time warping of human movement. The studied paradigm included the assessment of repetitive well-defined movements in different temporal distortion contexts. Repetitive motion is present in several circumstances such rehabilitation exercises, human gait dynamics, and movements that employees execute during the labor day in certain job activities.

In this subsection, we will present an experiment based upon time series retrieved using inertial sensors during the execution of repetitive motion. We created a dataset with a total of 240 signals retrieved by six different subjects that executed ten repetitions of a well-defined task under four distinct sets. The movements performed during each task consisted of: grasping a *solderless breadboard* used to build electronic circuits; placing the board on a defined position and welding a single perforation in each repetition; grasping the welded board and move it to a defined position. The difference among each class is based on the temporal criterion used by the subjects to perform the task as illustrated on Fig. 6.

Table 1
Results of Γ across the UCR archive. The SW-DTW was applied using $\alpha = 0.5$ and $\delta = 0.05 \times N$. In the DTW_R a 5% warping window was used. Higher values translate improper alignments and lower values represent correct alignments. The best alignment value for each dataset among the different DTW alignment variations is highlighted in bold font.

Dataset	DTW	DTW_R	DDTW	SW-DTW	Dataset	DTW	DTW_R	DDTW	SW-DTW
Adiac	0.900	0.743	1.528	**0.517**	Meat	1.640	1.164	0.930	**0.521**
ArrowHead	1.255	1.042	0.637	**0.457**	MedicalImages	1.405	**0.602**	0.694	0.654
Beef	1.484	1.247	0.684	**0.294**	MiddlePhalanxOutlineAgeGroup	0.821	0.607	0.676	**0.246**
BeetleFly	0.810	0.804	1.088	**0.196**	MiddlePhalanxOutlineCorrect	0.889	0.627	0.733	**0.307**
BirdChicken	1.125	1.084	0.640	**0.352**	MiddlePhalanxTW	0.817	0.607	0.738	**0.278**
Car	1.222	1.067	1.211	**0.568**	MoteStrain	1.007	0.595	0.532	**0.421**
CBF	0.608	0.552	0.021	**0.001**	NonInvasiveFatalECGThorax1	1.612	1.435	0.445	**0.284**
ChlorineConcentration	0.523	0.510	0.103	**0.025**	NonInvasiveFatalECGThorax2	1.664	1.362	0.544	**0.386**
CinCECGtorso	1.640	1.445	**0.232**	0.266	OliveOil	1.630	1.165	0.978	**0.378**
Coffee	1.459	1.176	0.541	**0.225**	OSULeaf	0.966	0.937	0.937	**0.270**
Computers	0.557	0.328	**0.000**	**0.000**	PhalangesOutlinesCorrect	0.770	0.586	0.713	**0.264**
CricketX	1.116	1.033	0.079	**0.019**	Phoneme	0.707	0.702	0.230	**0.006**
CricketY	1.206	1.078	0.097	**0.032**	Plane	0.827	0.776	0.824	**0.253**
CricketZ	1.162	1.018	0.080	**0.026**	ProximalPhalanxOutlineAgeGroup	0.816	0.608	0.800	**0.316**
DiatomSizeReduction	1.000	0.859	1.426	**0.551**	ProximalPhalanxOutlineCorrect	0.854	0.564	0.779	**0.309**
DistalPhalanxOutlineAgeGroup	0.698	0.544	0.616	**0.213**	ProximalPhalanxTW	0.803	0.628	0.855	**0.323**
DistalPhalanxOutlineCorrect	0.738	0.576	0.602	**0.226**	RefrigerationDevices	0.588	0.533	0.002	**0.000**
DistalPhalanxTW	0.757	0.581	0.717	**0.235**	ScreenType	0.546	0.301	0.001	**0.000**
Earthquakes	0.016	0.016	**0.000**	**0.000**	ShapeletSim	0.084	0.084	0.018	**0.000**
ECG200	0.893	0.658	**0.116**	0.162	ShapesAll	1.188	1.088	0.724	**0.425**
ECG5000	1.235	1.004	0.139	**0.117**	SmallKitchenAppliances	0.359	0.138	**0.000**	**0.000**
ECGFiveDays	0.947	0.758	0.153	**0.115**	SonyAIBORobotSurface1	0.292	0.242	0.052	**0.021**
ElectricDevices	0.210	0.096	0.007	**0.002**	SonyAIBORobotSurface2	0.120	0.109	0.048	**0.005**
FaceAll	0.382	0.373	0.229	**0.021**	Strawberry	1.525	1.177	0.757	**0.433**
FaceFour	0.567	0.559	0.006	**0.000**	SwedishLeaf	0.809	0.697	0.679	**0.325**
FacesUCR	0.369	0.365	0.216	**0.018**	Symbols	1.208	0.847	0.980	**0.613**
FiftyWords	1.276	0.988	1.081	**0.414**	SyntheticControl	0.233	0.202	0.030	**0.009**
Fish	1.171	0.996	1.755	**0.619**	ToeSegmentation1	1.249	1.107	0.207	**0.086**
FordA	0.683	0.683	0.625	**0.000**	ToeSegmentation2	1.399	1.180	0.321	**0.152**
FordB	0.658	0.658	0.580	**0.000**	Trace	1.926	1.066	**0.127**	0.292
GunPoint	1.750	0.643	1.443	**0.479**	TwoLeadECG	1.037	0.661	**0.443**	0.484
Ham	1.134	1.066	0.494	**0.042**	TwoPatterns	0.044	0.044	0.008	**0.000**
HandOutlines	0.983	0.959	1.214	**0.446**	UWaveGestureLibraryAll	1.314	1.218	0.517	**0.131**
Haptics	1.506	1.339	**0.363**	0.481	UWaveGestureLibraryX	1.197	0.861	0.784	**0.429**
Herring	1.210	1.119	1.112	**0.538**	UWaveGestureLibraryY	1.367	0.874	0.805	**0.573**
InlineSkate	1.786	1.493	**0.081**	0.575	UWaveGestureLibraryZ	1.330	0.893	0.938	**0.563**
InsectWingbeatSound	1.473	1.017	0.876	**0.290**	Wafer	1.479	0.530	0.990	**0.169**
ItalyPowerDemand	0.629	**0.000**	0.269	0.111	Wine	1.262	0.978	0.870	**0.591**
LargeKitchenAppliances	0.834	0.189	**0.001**	0.018	WordSynonyms	1.249	0.989	0.992	**0.416**
Lightning2	1.202	1.007	0.096	**0.010**	Worms	1.428	1.324	0.292	**0.135**
Lightning7	1.140	0.863	0.068	**0.023**	WormsTwoClass	1.562	1.478	0.244	**0.109**
Mallat	1.523	1.349	1.114	**0.324**	Yoga	1.148	1.053	1.167	**0.465**

Fig. 6. Summary of the proportional timings for task execution. The grasp movement is depicted by "G" and the soldering process is depicted by "S".

The subjects executed the repetitions under four distinct sets: *Normal*, where the subjects executed the movements at standard speed; *Short*, where the soldering process duration is approximately half of *Normal* speed; *Extended*, where the soldering process takes the double of the standard speed; and *OverallSlow*, where the subject completes the entire task taking approximately the double of time from standard speed.

Inertial information was retrieved using a custom IMU developed by Fraunhofer AICOS [31] streaming gyroscope data at 100 Hz. Since the device was placed on the wrist of each subject, it was possible to retrieve data that contained quasi-periodic sequences corresponding to all the task repetitions performed by the subject. The signals were manually segmented based on the beginning and end of each task and the SW-DTW and TAM were applied for each segmented window.

We divided our study in two perspectives: (1) using one of the subjects as a reference to establish a comparison between a reference subject and the group of the remaining subjects. The objective was to produce a representative distance of how similar the movements were performed against the reference; (2) provide a time series classification example, using TAM as a feature.

4.2.1. Reference movement comparison

In this approach we used a reference subject who recorded the reference movements that were compared against the group of remaining subjects. The reference was recorded by a single subject executing movements at standard and predefined timings which are described by the *Normal* set. Since the reference's model must at best represent the correct motion with inherent variability, we firstly interpolated all the tasks performed by the reference volunteer to the mean task length duration. Secondly, we computed the mean signal among all the interpolated signals. Since no previously signal alignment was performed, in case we had considered that the reference signal was the mean of the interpolated signals we were introducing artifacts during the process of mean calculation. In order to overcome this issue, we choose as reference task the one with minimum TAM value in comparison with the previously calculated mean. Therefore, the reference time series consists of the task repetition of the reference subject which potentially best

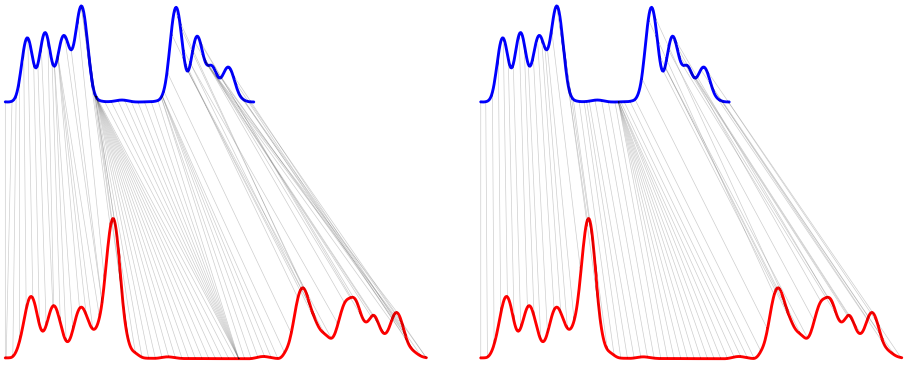

Fig. 7. Alignments between the reference time series (blue) and a repetition performed by a subject during the *OverallSlow* set (red). A comparison is provided between the DTW alignment (left) and the SW-DTW (right) with $\alpha = 0.05$ and $\delta = 2$ seconds. The interval where the grasp movement occurs is depicted by "G" and the interval where the soldering process is executed is depicted by "S". For presentation purposes the alignment lines are not displayed for the entire set of samples.

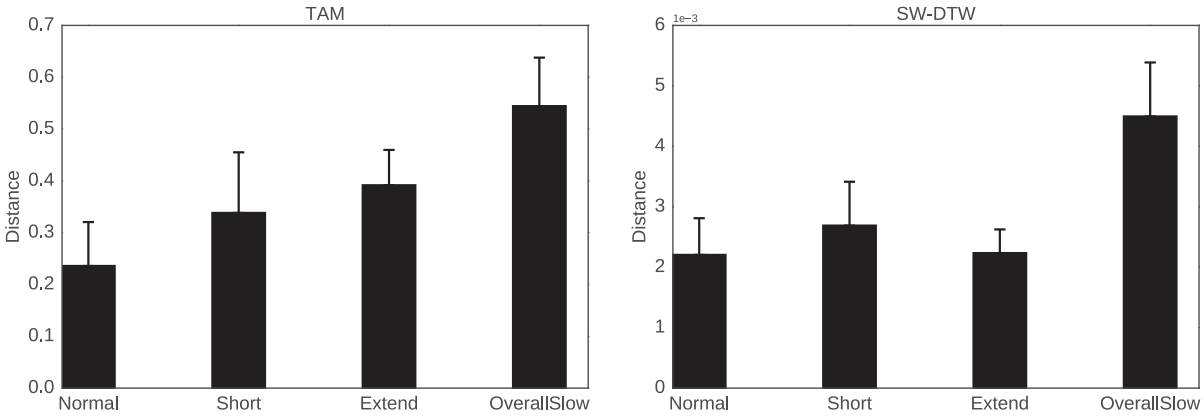

Fig. 8. Mean and standard deviation SW-DTW and TAM values of all subjects in different set speeds.

corresponds to the minimum temporal misalignment in comparison with its own mean.

Fig. 7 illustrates an example of the alignment established by the SW-DTW, where the reference time series is compared against a signal acquired from another subject executing the *OverallSlow* set. The signals comprise gyroscope filtered data and the prominent events correspond to the executed movements necessary to achieve the task. The plateau on both series corresponds to the moment where the subject is actually placing the iron tip against the perforation to accomplish the soldering.

We can observe a misadjustment between peaks corresponding to the same event. In the *OverallSlow* example the peaks occur in different instants and they tend to show a temporal offset to the right. Therefore, we can declare that they are temporally delayed relative to the reference time signal. In line with the results from 4.1, the visual comparison potentially suggests that the alignment produced by SW-DTW reduces the singularity issues and allows a more accurate TAM calculation in comparison with the DTW methodology. In fact, even in the segment which is prone to lead to singularities, such the plateau, the SW-DTW seems to reasonably map the delay among the two series, which is not observed in the DTW as both an advance and delay are present since two singularities occur.

After manually segment all tasks, the distances between the reference time series and the remaining signals of the dataset acquired in the four contexts were calculated using the SW-DTW and the TAM. Fig. 8 summarizes the results of the mean and standard deviation values for the SW-DTW and TAM distances between the reference and the group of the remaining subjects for each set.

Since the signals are similar-alike in amplitude, the SW-DTW values are similar between *Normal*, *Short*, and *Extend* sets. The *OverallSlow* set produced an higher score since angular acceleration may become attenuated when the subject tries to execute the task at a slower pace. The analysis using TAM shows a similar pattern with an exception of the decrease relative distance of the *Extended* set. The highest similarity between sets is present between the *Short* and *Extend*, despite the fact they still continue to exhibit higher values in comparison to the *Normal*. This result can be explained since the TAM measures the overall time warping between series. Since the ratio of advance in the *Short* set is similar to the ratio of delay in the *Extend* set, they end up showing the same extent of overall warping.

The advantage of using TAM to complement the analysis lies on the fact we are still able to retrieve further information if we examine the ratios of delay and advance for each set. Since the *Short* set comprises an advance it is expected that $\overrightarrow{\psi} > \overleftarrow{\psi}$. On the other hand, as *Extend* constitutes a delay, one can anticipate $\overrightarrow{\psi} < \overleftarrow{\psi}$. Those assumptions are supported by the results outlined in Fig. 9.

The ratio between both parameters in the *Normal* suggests that although the subjects try to follow the predefined reference timings for movement execution, there is an inherent variability associated with the movements required to complete the task. On the other side, the *Short* set ratio possesses a significantly higher weight for advance, in contrast with the *Extend*, which denotes a predominant weight of delay, as expected by the nature of how the task was performed in each respective set. *OverallSlow* shows the more significant weight increase for delay.

This study demonstrated the potential of TAM to discriminate between different time warping contexts of the same activity. The

Table 2
Accuracy (mean ± standard deviation) after k-fold cross validation in comparing distinct distance functions. A total of 20 folds was evaluated. The best value among the different classifiers is highlighted in bold font.

Measurement	T	MSE	DTW	SW-DTW	TAM(DTW)	TAM(SW-DTW)
Accuracy ($\mu \pm \sigma$)	(0.85 ± 0.05)	(0.80 ± 0.04)	(0.90 ± 0.03)	(0.92 ± 0.03)	(0.37 ± 0.04)	$\mathbf{(0.96 \pm 0.03)}$

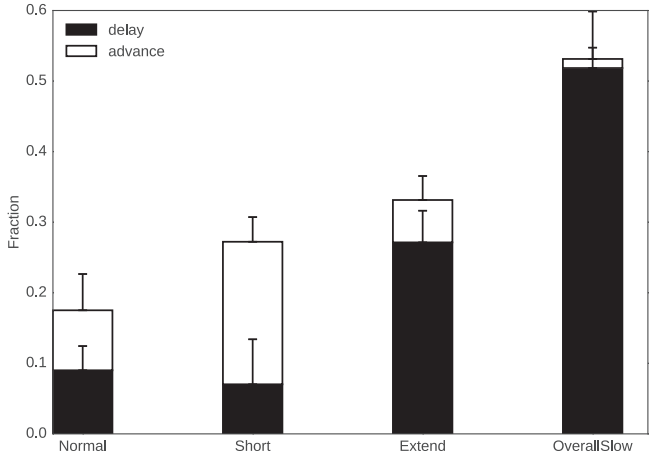

Fig. 9. Mean and standard deviation values for $\overrightarrow{\psi}$ and $\overleftarrow{\psi}$ for all subjects in different set activities.

movements performed in all sets to accomplish the task were exactly the same. However, the timings of the movements executed were different and lead to non linear warping between time series, which were successfully described by our proposed distance. The analysis using TAM is more informative than SW-DTW as it is potentially able to discriminate the results from *Short* and *Extended*, which shown similar distance values using the SW-DTW.

In this approach we were able to demonstrate the ability to produce meaningful scores using TAM in a scenario where a group of subjects were compared against a previously recorded reference. In the next subsection we will present a classification example using our proposed measurement.

4.2.2. Time series classification

In this approach we merged the data from all subjects into 4 classes that correspond to each set (*Normal, Short, Extended*, and *OverallSlow*). We implemented a 1-NN classifier using the following distance functions: the absolute distance between the length of two series T; Mean Square Error (MSE), which corresponds to the Euclidean distance after interpolating both sequences to the length of the longest sequence; DTW; SW-DTW with $\alpha = 0.25$ and $\delta = 4s$; TAM(DTW), which corresponds to the TAM value calculated using the DTW alignment and TAM(SW-DTW), which corresponds to the TAM value calculated using the SW-DTW alignment.

We divided the subjects into training and testing sets with an equal number of elements and used a k-fold cross validation to cover the entire group of possibilities between the distribution of the subjects in the training and testing sets. The results are presented in Table 2.

The highest accuracy was achieved by the TAM calculated using the SW-DTW alignment. However, a significant difference is present when using the DTW alignment. These results reinforce the fact that the quality of the previous time series alignment is crucial to achieve representative performance when using our proposed measurement. The DTW is prone to produce singularities more often than SW-DTW and they will tend to increase the variability of the distance between time series of the same class which downgrades the classifier's performance. We can also observe a performance increase when using TAM in comparison with T. This fact suggests that for a given set, there is inherent variability among subjects associated to the total time spent in each task. The TAM performs a description during the complete signal length and, consequently, detects the temporal pattern associated for each set and not strictly the difference between the signal lengths.

On the amplitude domain, we were also able to achieve good performance. The similarity between the DTW and SW-DTW accuracies suggest that even with singularities, the DTW was able to distinguish the 4 classes. It is worth to mention that DTW and TAM measure different realities (amplitude and time, respectively). This fact explains how both DTW and SW-DTW achieved good performance while a different behaviour was present in the TAM calculation based on DTW and SW-DTW alignments, since there was a significant difference between accuracies calculated. Finally, the MSE achieved the lower performance on the amplitude based features as the linear interpolation between the sequences is not able to model the non linear warping that is present between different sets. This non linear warping is correctly modelled using DTW based algorithms as they produce a more intuitive distance as discussed in subsection 2.1.

5. Conclusions and Future Work

One of the most important topics in the context of the study of time series is the development of novel measures able to characterize each signal. As most of the distances are based in amplitude, a true time distance able to characterize the degree of time warping between two sequences can be useful in a wide range of domains. We believe that this distance can be applied in several contexts, such as human movement analysis and electrophysiological data.

This paper presents two relevant contributions to the aforementioned domain. The alignment between time series can provide a pairwise relationship between elements which translates information on the time domain. One of the state-of-the-art techniques to perform the alignment is based on DTW algorithm. There are however some circumstances, arising when the algorithm tries to express amplitude variability in the Y-axis by improper warping the X-axis, where incorrect alignments may eventually be present. The distance function used in the DTW only uses a point-to-point distance and does not assess the context where a particular time instant is inserted. One of our contributions was the development of a new local cost distance for the DTW algorithm. Using a window instead of an element-to-element approach potentially allows to prevent singularities by looking in a region which takes into account a weighting between the amplitude and the first discrete derivative of both signals. Despite the achieved results showed a significant decrease of singularities, our approach is computationally demanding and may require detailed optimization for real-time usage.

The major contribution of our work is a comprehensive technique to characterize the degree of warping between time series. In this paper we started by presenting a detailed analysis of DTW optimal warping path. The vertical, horizontal and diagonal segments can deliver information related to the delay, advance and phase, respectively, between two given time series. Our approach tries to measure the cost that one sequence must perform to match the temporal requirements of a given reference se-

quence. A limitation of the TAM is that relies on the alignment quality between two series in order to use the optimal alignment path. Therefore, an improved alignment was achieved using the SW-DTW approach. The TAM distance was successfully applied to both artificial and real time series data. We demonstrated two examples of applicability to this novel measurement: the TAM can be used as quality index to establish an comparison between different signal alignment methodologies. Our results show that SW-DTW is prone to reduce singularities in comparison with the evaluated alternatives; we also demonstrated the possibility to discriminate time warping differences of human repetitive movement. In this case, although time series demonstrated amplitude similarity as the movements being executed were equal, they show different degrees of non-linear warping according to the nature of each set.

It is important to emphasize that in this paper we do not intend to achieve a generic higher performance of TAM in comparison with DTW. The two approaches measure different realities: SW-DTW measures distance in the amplitude domain and it will be more suitable for classification in datasets with amplitude variability among classes; TAM express distance in the temporal domain and it is more suitable to classify datasets with minor amplitude deviation and high temporal variability. Being two measurements of different nature they will be applied according to different realities.

The results obtained also suggested that TAM is sensible to the previous pairwise alignment and, therefore, a correct adjustment of the α and δ parameters is required when using SW-DTW. While we based our study in empirical derivation of the best values for each parameter, in future it will be required a detailed analysis of their influence in the alignment quality and ultimately in the TAM calculation. In the present work we analysed the optimal warping path from a discrete perspective. By using the pairwise alignment provided by DTW, the warping path segments follow discrete slopes. The future work of this approach will consist in generalizing the TAM calculation from the optimal warping path to the continuous domain.

Acknowledgement

This work was supported by North Portugal Regional Operational Programme (NORTE 2020), Portugal 2020 and the European Regional Development Fund (ERDF) from European Union through the project Symbiotic technology for societal efficiency gains: Deus ex Machina (DEM) [NORTE-01-0145-FEDER-000026].

References

[1] E. Keogh, C.A. Ratanamahatana, Exact indexing of dynamic time warping, Knowledge and Information Systems 7 (3) (2005) 358–386, doi:10.1007/s10115-004-0154-9.
[2] C. Ratanamahatana, E. Keogh, Making time-series classification more accurate using learned constraints, in: Proc. of Sdm Int'L Conf, 2004, pp. 11–22. 10.1.1.215.1648
[3] C.A. Ratanamahatana, J. Lin, D. Gunopulos, E. Keogh, M. Vlachos, G. Das, Mining time series data, in: Data Mining and Knowledge Discovery Handbook, 2010, pp. 1049–1077, doi:10.1007/978-0-387-09823-4_56.
[4] T. Araújo, N. Nunes, H. Gamboa, A. Fred, Generic Biometry Algorithm Based on Signal Morphology Information: Application in the Electrocardiogram Signal, Springer International Publishing, Cham, 2015, pp. 301–310, doi:10.1007/978-3-319-12610-4_19.
[5] J. Barth, C. Oberndorfer, C. Pasluosta, S. Schlein, H. Gassner, S. Reinfelder, P. Kugler, D. Schuldhaus, J. Winkler, J. Klucken, B.M. Eskofier, Stride segmentation during free walk movements using multi-dimensional subsequence dynamic time warping on inertial sensor data, Sensors 15 (3) (2015) 6419, doi:10.3390/s150306419.
[6] G.E. Batista, X. Wang, E.J. Keogh, A complexity-invariant distance measure for time series, in: SDM, Vol. 11, SIAM, 2011, pp. 699–710, doi:10.1137/1.9781611972818.60.
[7] J. Aach, G.M. Church, Aligning gene expression time series with time warping algorithms, Bioinformatics 17 (6) (2001) 495–508, doi:10.1093/bioinformatics/17.6.495.
[8] D. Clifford, G. Stone, I. Montoliu, S. Rezzi, F.P. Martin, P. Guy, S. Bruce, S. Kochhar, Alignment using variable penalty dynamic time warping, Analytical Chemistry 81 (3) (2009) 1000–1007, doi:10.1021/ac802041e.
[9] I.P. Machado, A.L. Gomes, H. Gamboa, V. Paixo, R.M. Costa, Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization, Information Processing and Management 51 (2) (2015) 204–214, doi:10.1016/j.ipm.2014.07.008.
[10] X. Xia, X. Song, F. Luan, J. Zheng, Z. Chen, X. Ma, Discriminative feature selection for on-line signature verification, Pattern Recognition 74 (2018) 422–433.
[11] S. Chu, E.J. Keogh, D.M. Hart, M.J. Pazzani, Iterative deepening dynamic time warping for time series, in: SDM, SIAM, 2002, pp. 195–212, doi:10.1137/1.9781611972726.12.
[12] M. Müller, Dynamic time warping, Information retrieval for music and motion (2007) 69–84, doi:10.1007/978-3-540-74048-3_4.
[13] M.E. Munich, P. Perona, Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification, in: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, Vol. 1, IEEE, 1999, pp. 108–115, doi:10.1109/ICCV.1999.791205.
[14] P.H. Eilers, Parametric time warping, Analytical chemistry 76 (2) (2004) 404–411, doi:10.1021/ac034800e.
[15] Y. Chen, B. Hu, E. Keogh, G.E. Batista, Dtw-d: time series semi-supervised learning from a single example, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2013, pp. 383–391.
[16] L. Chen, R. Ng, On the marriage of lp-norms and edit distance, in: Proceedings of the Thirtieth International Conference on Very Large Data Bases - Volume 30, VLDB '04, VLDB Endowment, 2004, pp. 792–803.
[17] L. Chen, M.T. Özsu, V. Oria, Robust and fast similarity search for moving object trajectories, in: Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, SIGMOD '05, ACM, New York, NY, USA, 2005, pp. 491–502, doi:10.1145/1066157.1066213.
[18] P.F. Marteau, Time warp edit distance with stiffness adjustment for time series matching, Pattern Analysis and Machine Intelligence, IEEE Transactions on 31 (2) (2009) 306–318, doi:10.1109/TPAMI.2008.76.
[19] F. Petitjean, A. Ketterlin, P. Gançarski, A global averaging method for dynamic time warping, with applications to clustering, Pattern Recognition 44 (3) (2011) 678–693.
[20] M. Morel, C. Achard, R. Kulpa, S. Dubuisson, Time-series averaging using constrained dynamic time warping with tolerance, Pattern Recognition 74 (2018) 77–89.
[21] J. Lines, A. Bagnall, Time series classification with ensembles of elastic distance measures, Data Mining and Knowledge Discovery 29 (3) (2014) 565–592, doi:10.1007/s10618-014-0361-2.
[22] U. Mori, A. Mendiburu, J. Lozano, TSdist: Distance measures for time series data, in: R Foundation for Statistical Computing, 2016. http://cran.xl-mirror.nl/web/packages/TSdist/index.html.
[23] M. Barandas, H. Gamboa, J. Fonseca, A real time biofeedback system using visual user interface for physical rehabilitation, Procedia Manufacturing 3 (2015) 823–828, doi:10.1016/j.promfg.2015.07.337.
[24] E.J. Keogh, M.J. Pazzani, Derivative dynamic time warping, SIAM, 2001, doi:10.1137/1.9781611972719.1.
[25] Y. Chen, E. Keogh, B. Hu, N. Begum, A. Bagnall, A. Mueen, G. Batista, The UCR time series classification archive (july 2015).
[26] R. Muscillo, S. Conforto, M. Schmid, P. Caselli, T. D'Alessio, Classification of motor activities through derivative dynamic time warping applied on accelerometer data, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 4930–4933, doi:10.1109/IEMBS.2007.4353446.
[27] A. Zifan, S. Saberi, M.H. Moradi, F. Towhidkhah, Automated ECG segmentation using piecewise derivative dynamic time warping, Int. J. Biol. Med. Sci 1 (2006) 181–185.
[28] Y.S. Jeong, M.K. Jeong, O.A. Omitaomu, Weighted dynamic time warping for time series classification, Pattern Recognition 44 (9) (2011) 2231–2240, doi:10.1016/j.patcog.2010.09.022.
[29] Z. Zhang, P. Tang, R. Duan, Dynamic time warping under pointwise shape context, Information Sciences 315 (2015) 88–101, doi:10.1016/j.ins.2015.04.007.
[30] J. Zhao, Z. Xi, L. Itti, metricdtw: local distance metric learning in dynamic time warping, ArXiv preprint arXiv:1606.03628.
[31] Fraunhofer AICOS, White paper: A day with pandlets, tech. rep., Research Center for Assistive Information and Communication Solutions, 2016.

Duarte Folgado received his MSc in Biomedical Engineering from the Faculty of Sciences and Technology of NOVA University of Lisbon. After finishing the MSc he continued working as a scientist at Fraunhofer AICOS. His main research interests include computer science techniques, signal processing, and embedded systems for Assistive Environments.

Marília Barandas received her MSc in Biomedical Engineering from the Faculty of Sciences and Technology of NOVA University of Lisbon. After completing her master's thesis, Marília was invited to lecture Medical Information Systems at FCT-UNL and to join Centre of Technology Systems, Portugal. Since April 2015, she is a scientist at Fraunhofer AICOS, focusing in indoor locations solutions based on smartphones' built-in inertial sensors.

Ricardo Matias is a Researcher at Champalimaud Centre for the Unknown of the Champalimaud Foundation. He has a Ph.D in Human Kinetics from the University of Lisbon. His research combines computational biomechanics with machine learning to help uncover the mechanisms that trigger the decline from healthy mobility to movement pathology.

Rodrigo Martins is an Assistant Professor at the Physiotherapy Department of the Portuguease Red Cross Health School. He is a PhD candidate in Biomechanics at Human Motricity Faculty, University of Lisbon, Neuromechanics Research Group Interdisciplinary Centre for the Study of Human Performance (CIPER), currently working in human motion pattern recognition mainly gait.

Miguel Carvalho is a scientist, researcher, and professor of textile engineering at University of Minho. Degree in Textile Engineering, MSc in Design and Marketing, PhD in Textile Engineering â Clothing Technology. Focus: clothing and textile design, ergonomics, anthropometrics, development of functional and interactive materials, production planning and control, work study, teamwork.

Hugo Gamboa received his PhD in Electrical Engineering from Instituto Superior Técnico, Universidade de Lisboa, and he is an Assistant Professor in the Physics Department of Faculty of Sciences and Technology of NOVA University of Lisbon, Portugal. He is also a Senior Scientist at Fraunhofer Portugal AICOS. He has authored more than 100 papers in conferences and journals. His research activities focus on biomedical instrumentation and biosignals processing and classification.