Paired tenzotremorogramms structure similarity analysis based on time series distance functions: problem formulation

O.V. Zhvalevsky
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences
St. Petersburg, 14-th Linia 39, St. Petersburg, 199178, Russia
ozh@spiiras.ru

Abstract. The paper is concerned with paired tenzotremorogramms that are a result of finger force registration (for both human arms) via sensitive platform with tenzosensor. The main hypothesis is: the dynamics of force hold process differs for different patient category (for healthy patients and for patients who have some pathology). In the paper some approaches to paired tenzotremorogramms structure similarity analysis are investigated. These approaches are based on application of time series distance functions.

1. Introduction
Parkinson’s Disease diagnostics requires high quality experimental data. Tremor is the most popular source of experimental data [1]. There are many methods for tremor registration, but not all of them provide exact and reliable information about movement control system (movement construction system according to Bernstein [2]). The most of these methods deal with voluntary movements of extremities. Each registration method produces tremorogramms. Tenzotremorogramms (TTG) are the result of finger force registration (for both human arms) via sensitive platform with tenzosensor [3]. But this kind of experimental data contain information about unvoluntary movements of extremities (human arms). Measurements are performed in isometric mode of tremor registration, which makes TTG a unique type of time series [4]. So, this kind of experimental data require specific registration technique and data processing technique.

Original TTG registration technique including a biofeedback is as follows: each sensitive platform is linked with graphic label on computer screen (one label corresponds to left arm, another label corresponds to right arm), and examinee should align these two labels during determined time and hold determined level of the force. TTG registration technique generates complex force hold process. The main hypothesis is: the dynamics of force hold process differs for different patient categories (for healthy patients and for patients who have some pathology). Left TTG (TTG corresponding to left arm) and right TTG (TTG corresponding to right arm) are called (in this paper) paired TTGs, and they are considered (in this paper) as primary exploration subject. The main question is: how can we evaluate dynamics of force hold process? In this paper we consider some approaches to paired TTGs structure similarity analysis. These approaches are based on application of time series distance functions.

The rest of the paper is organized as follows. Section 2 reports some preliminary notes, which precede the main part of the paper. In section 3 the main problem is formulated. Section 4 is devoted...
to data processing technique representation. Also, there are some conclusions and directions of future research work in section 5.

2. Preliminaries

There are some basic notions, which precede main problem formulation and data processing techniques representation. The main notion is time series distance function [5]. Generally, distance functions are used for time series clustering and classification. The whole time series clustering is not so useful for paired TTGs analysis, but, at the same time, time series fragments clustering is useful for paired TTGs analysis, but it requires accurate and effectiveness approach to time series fragments recognition. This approach is obviously based on design of time series distance profile and distance matrix. Time series profile design provides experimental data indexing, where each given time series is represented as a sequence of patterns or symbols. Distance matrix design leads to some more sophisticated distance functions, and as it is showed in the paper, leads to new approach to time series comparison. At last, time series distance values calculation requires (in some important cases) some specific time series transformations.

2.1. Time Series Distance Functions

Distance function is a nonnegative real function that estimates dissimilarity of two given objects. Time series distance function is a distance function, which is used for time series. In common case, if \( d \) is a selected distance function, then a distance between two given time series \( X \) and \( Y \) is denoted as \( d(X, Y) \). Distance value is small for similar objects, and, otherwise, is large for dissimilar objects. Two given time series can be similar in one aspect of time series analysis, but they can be dissimilar in the other aspect. Therefore, we should use different time series distance functions to manage several aspects of time series analysis. Researchers try to propose the best function for solving their problems, but not all of them are unified or multi-purpose time series distance functions.

Time series distance functions can be distinctly split into two principal groups. The first group is formed by distance functions, which deal with raw time series data. Lock-step distance functions represent the first class of time series distance functions. These time series distance functions directly use point-to-point comparisons, when one point of one time series is compared with one point of another time series. Euclidean distance and another functions based on Lp-norms (e.g. Manhattan Distance and Maximum Distance) are the most popular examples of these distance functions [6, 7]. Elastic functions, which allow one-to-many or one-to-none points comparisons, represent the second class of time series distance functions. The main example of them is DTW (dynamic time warping) [8]. Some of elastic distance functions are based on edit distance for strings, e.g. EDR (edit distance on real sequence) [9], ERP (edit distance with real penalty) [10] and Swale (sequence weighted alignment model) [11]. LCSS (Longest Common SubSequence) is an example of distance function, which allows one-to-many points comparison and is based on edit distance [12].

The second distance function’s group consists of distance functions, which require special defined time series aggregations and representations. SAX (Symbolic ApproXimation) of time series (based on Piecewise Aggregate Approximation and time series values symbol map) leads to MINDIST [13]. Time series transformation to a sequence of threshold-crossing intervals, where the points within each time interval have a value greater than given threshold, results in TQuEST distance [14]. Finally, SpADe is an example of distance function based on time series patterns, which allow shifting and scaling in both the temporal and amplitude dimensions [15].

2.2. Time Series Distance Profile

Distance profile is a sequence of distance values calculated for the window sliding along one time series and fixed time series fragment. Distance profile can be interpreted as a special time series representation. Distance profile building is needed for time series indexing, clustering and classification.
Let $S$ be a time series (search area) and $P$ be some pattern (or query example). Also, let $X$ be a segment of time series $S$. At first, we should select some time series distance function $d$ and threshold $\varepsilon$, so we can calculate distance value $d(P, X)$ and compare it with the given threshold $\varepsilon$. If distance value $d(P, X)$ is smaller than $\varepsilon$, time series segment $X$ is called occurrence of the given pattern $P$ in time series $S$. Threshold fixation determines all occurrences of the given pattern $P$ to time series $S$. Then, we should consider a sequence of time series segments or fragments, and calculate distance value $d(P, X)$ between the given pattern $P$ and time series segment $X$ for each segment $X$ from the sequence. As a result, a function $d(P; S)$ built for time series $S$ and the given pattern $P$ is called distance profile. The building of distance profile explicitly suggests the same procedure of sliding window (having a particular width), which is moving along time series $S$. Each position of sliding window corresponds to time series segment of the given time series.

In general, we can select a set of certain patterns (time series segment types, events or symbols) and represent each time series (from such experimental data base) in symbolic representation form.

### 2.3. Time Series Distance Matrix

Distance matrix is a more complicated notion than distance profile.

Let $\left[\begin{array}{c}X_i \\ Y_j \end{array}\right]$ be a segmentation of the first time series $X$, and $\left[\begin{array}{c}X_i \\ Y_j \end{array}\right]$ be the same segmentation of the second time series $Y$. So, we can build a matrix, so-called distance matrix, for two given time series, where each element is a distance between segment $X_i$ of the first time series $X$ and segment $Y_j$ of the second time series $Y$ ($j = 1, K$).

It is clearly noticed that each row or each column of the distance matrix is a distance profile of one time series built for separate segment of another time series. Thus, distance matrix contains full information for comparative analysis of two given time series in relation to fixed length of segments and fixed distance function. Furthermore, it provides some more specific time series distance functions allowing more elastic and deeper approach to time series analysis.

Fixed length of segments is the main disadvantage of distance matrix. Optimal segment length selection method does not exist. Moreover, sequence of distance matrixes corresponding to a set of segment lengths contains more complete information about analyzed time series. At the same time, hierarchy segmentation is more accurate and adequate technique for time series structure analysis. Also, we can expect that different distance functions may be used at different levels of hierarchy so the main challenge of time series analysis is to find relation between differ levels of hierarchy. As a result, we could provide an “effectiveness and efficiency” time series distance function for certain research domain.

### 2.4. Time series transformations

There exist some time series transformations usually preceding distance function application. Particularly, these transformations can be an essential part of some distance function definitions in several specific cases.

Firstly, time series distance function between two given time series $X$ and $Y$, denoted as $d(X, Y)$, can be defined as the same distance between their transformations denoted as $d(\tilde{X}, \tilde{Y})$, where $\tilde{X}$ and $\tilde{Y}$ are the transformed time series or time series representations for time series $X$ and $Y$, accordingly.

Secondly, transformation can be an inherent part of distance function definition. Let $D(x, y)$ be a basic (or intrinsic) distance function between two given objects $x$ and $y$; also, $x(\tau) = \xi(\tau; X)$ and
\( y(\tau) = \eta(\tau; Y) \), accordingly. So, let \( d(X, Y) \) be an extremum value of \( D[x(\tau), y(\tau)] \), minimum or maximum value depending of task in consideration.

The first important class of time series transformations is **normalization**. Normalized time series have values between 0 and 1. It’s useful for development of unified scale for correct time series comparative analysis. Also, we can use **standardization** for the same task.

The second important class of time series transformations is **aggregation**. For example, Piecewise Aggregate Approximation (PAA) of time series \( X \) is a time series \( \overline{X} \), where each value is the mean value for certain segment. Let \( \left[ X_{\Delta \tau} \right]_{k=1}^{K} \) be a segmentation of the time series \( X \). Also, let \( \overline{X}_k \) be a time series segment represented by constant value \( \overline{x}_k = \text{mean}(X_k) \). So, \( \overline{X} = \left[ \overline{X}_k \right]_{k=1}^{K} \) is PAA of time series \( X \). In many applications PAA is used as basic time series representation for building more meaningful time series representations (i.e., as the first step of some specific algorithms). In common case, we can apply more complex piecewise aggregation (not only mean values), linear and non-linear models for representation of time series on each separate segment.

### 3. Problem formulation

Paired TTGs provide new level of time series analysis under investigation.

Indeed, let \( d(X, Y) \) be a distance between two given time series \( X \) and \( Y \), where \( d \) is a selected distance function. In context of time series clustering, the main goal of experimental data processing is to group selected objects in separate groups or clusters in accordance with given distance function. Each cluster is ideally a compact group of points in feature space. Optimal distance function selection determines accurate time series clustering.

In opposite to it, let \( X \) be a pair of two time series denoted as \( X = [X_l, X_r] \), where \( X_l \) and \( X_r \) are the left TTG and the right TTG for object \( X \), accordingly; also let \( Y \) be a pair of two time series denoted as \( Y = [Y_l, Y_r] \), where \( Y_l \) and \( Y_r \) are the left TTG and the right TTG for object \( Y \), accordingly. So, we can calculate the two different internal distance: \( d(X_l, X_r) \) and \( d(Y_l, Y_r) \), accordingly. The main hypothesis is: \( d(X_l, X_r) \) is always smaller than \( d(Y_l, Y_r) \), when object \( X \) belongs to the first class and object \( Y \) belongs to the second class. Optimal time series distance function selection leads to optimal separation of selected objects into two classes: one class is (for example) “healthy” examinee group, and another class is the group of patients having some pathology (PD, for example).

There are some important aspects dictated by experimental data nature and structure. The first one is TTG classification, and the second one is TTG types classification.

#### 3.1. TTG classes

Which classes of TTG we should consider to solve the main problem?

The first class is a group of healthy people. These people haven’t any pathology, but we should take in account that some of them can be the patients staying in the nonmanifest stage of disease (PD, for example). The second class is a group of patients with PD, and the third class is a group of patients with parkinsonism (i.e., parkinsonian syndrome). Effective diagnostics must provide a good separation between two such classes. At last, the fourth class is a group of patients, who have some pathology not linked with PD and syndrome of parkinsonism.

These classes are indeed the patient categories, but real classes of TTG are the results of measurements carried out in certain measurement conditions. For example, PD-patient usually gets certain treatment, so there are two clearly separate TTG classes: the first one is the pure (background) TTG, which corresponds to measurement without treatment, and the second one is the ‘dirty’ (control) TTG, which corresponds to measurement with usual treatment.
3.2. TTG types
There are four types of TTG. Each TTG type corresponds to certain measurement method.

Measurement experiment consists of two principal parts. In the first part of experiment the patient uses fingers for holding the given force level only. The second part of experiment the patient uses whole arms for holding the given force level. Each part of experiment consists of two functional tests: the first one is the test for minimum force level, and the second one is the test for maximum force level. So, we have four different functional tests for each measurement method.

Let $X$ be experimental data (a set of time series), collected for the certain patient in separate time. Thus, we have four data blocks (TTG for each functional test), denoted as $X = [X_1, X_2, X_3, X_4]$ or as $X = [X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2}]$, where the first index corresponds to the part of experiment, and the second index corresponds to a number of functional test.

4. Data processing technique
There are traditionally two different approaches to time series analysis.

The first approach to time series analysis is based on time series characteristics calculations (via time series processing methods and algorithms) and feature extraction. Easy realization of this approach is to find time series characteristic, which provides easy decision rule formulation for given objects separation into desired classes. In common case, it is need to calculate a set of time series characteristics, and, hence, extract features, and design a system of decision rules based on extracted features.

The second approach to time series analysis is based on design of time series distance function and time series clustering.

There are three different types of clustering. The first one is whole time series clustering. There are several techniques proposed to increase “effectiveness and efficiency” of these clustering methods including dimension reduction and lower bounding calculation. The second one is the time series subsequence clustering, but there exist some aspects, which make this type of clustering meaningless. Indeed, the width of window sliding along time series is a result of voluntary selection, and there are no suggestions to find the best width. The third type of clustering is time series structure elements (patterns, events) clustering. This type of clustering is the updated and advanced version of the second type of clustering. Patterns are not given before clustering, but they are the results of it.

Experimental data structure dictated the using of certain data processing methods and algorithms. Let us to consider some aspects of data processing technique in detail.

4.1. Data Structure
Let $X_{i \in [1, N]}$ be a data sample (or selection), where each sample is a time series pair $X_i = [X_i^l, X_i^r]$ (where $i \in [1, N]$). So, we can build the column-vector of distance values $D = (d_i)_{i=1}^{N}$. Each element of this vector is the distance value $d_i = d(X_i^l, X_i^r)$ calculated between left TTG and right TTG ($i \in [1, N]$). Four TTG types produce matrix $D = [D_1, D_2, D_3, D_4]$, where each column corresponds to the certain TTG type.

4.2. Complex data processing
There are three different approaches to analysis of paired TTGs.

The first approach is to analyze each TTG type separately. Four TTG types produce four different analysis variants for each matrix $D_k$ ($k \in [1, 4]$). Each analysis variant is a “vertical” data analysis kind, where we are interested in data sample separation by particular time series distance function.
The second approach is to analyze the relation between different TTG types. This approach is a “horizontal” data analysis kind.

The third approach is to analyze all TTG types in complex. This approach is represented by complex data processing methods, which are taking into account both analysis kinds: “vertical” and “horizontal”.

4.3. TTG segmentation

Some time series processing methods require time series segmentation. Time series segmentation leads to dimension reduction. A more interesting and important challenge is to recognize the time series structure. The aim of recognition is to find similar fragments in paired TTGs, which reflect the process of force holding. Time series segmentation provides building many time series profiles for each segmentation under consideration. Moreover, hierarchy segmentation can help to overcome meaningless time series subsequence clustering, but it makes time series analysis more complex and more difficult to validate.

Let TTG be a time series with length \( N = 3000 \), which corresponds to 30 seconds (for samples frequency \( f_s = 100 \) Hz). There are 32 different acceptable segmentations, since \( 3000 = 2^3 \cdot 3^1 \cdot 5^3 \). So, we can build many acceptable hierarchy TTG segmentations.

4.4. TTG transformation

There is certain time series transformation in order to apply selected time series distance function. This transformation is organized as follows.

Let \( X \) be a normalized time series. Also, let \( \overline{X} \) be a PAA representation of time series \( X \) for certain segmentation. Distance between \( X \) and its PAA representation denoted by \( \hat{X} = X - \overline{X} \) estimates the accuracy of approximating long segments leading to a short and vary coarse approximation. Otherwise, short segments lead to a long or vary detailed approximation. So, we can select desired approximation quality and use it for distance value calculation for two given time series. Thus, we should calculate the distance \( d(\hat{X}, \hat{Y}) \) instead of the distance \( d(X, Y) \).

The full time series analysis is the calculation of a series of distances \( d(\hat{X}, \hat{Y}) \) for each segmentation from the given set of segmentations instead of calculation of the separate distance value \( d(X, Y) \).

4.5. Generalization of dynamic time warping

Which algorithm we should construct to solve the problem formulated in the paper? It’s obviously should be a generalized version of DTW (Generalized DTW) in the sense described as follows.

Let \( \left[ X_{k \leq i < N} \right] \) be a segmentation of the first time series \( X \), and \( \left[ Y_{k \leq i < N} \right] \) be the same segmentation of the second time series \( Y \). Also, let \( X_1 \) be the first segment of the first time series \( X \). Thus, we should find in the second time series \( Y \) a segment \( Y_j \), which is the best correspondence for segment \( X_1 \), and use this segment to find the best correspondence for it in the first time series \( X \), and so on. If we would act as described before we will get a sequence of segments belonging to two given time series \( X \) and \( Y \) (odd segments belong to the first time series \( X \), even segments belong to the second time series \( Y \)). On the other hand, we will get a sequence of time segment pairs (the first time series segment belongs to the first time series, and the second time series segment belongs to the first time series), so optimal path is the continual set of segment pairs.

Building optimal path corresponds to force holding process reconstruction. Hence, optimal time series distance function selection predetermines optimality of the built path. Standard DTW deals with separate points, but Generalized DTW deals with whole segments. These segments can have different length, and the segment’s sequence built for each time series becomes an important time series
representation form. So, the main problem is to find a relation between this representation form and patients category.

5. Conclusions
The paper is devoted to a new research area. The main problem is to develop the methods for paired TTG analysis. The paper is concerned with theoretical and methodological aspects of the problem. In the paper some approaches to paired TTG structure similarity analysis are investigated. These approaches are based on application of time series distance functions. It is showed that these approaches are usable for solving the main problem. Also, data processing technique is proposed to realize these approaches. Certain data structure, hierarchy segmentation, special time series transformation are the principal parts of this technique. Consideration of data processing technique leads to a challenge, that is to build a generalization of DTW (Generalized DTW) based on time series segments instead of time series points (for standard DTW). Designing accurate and robust algorithm for Generalized DTW is the first direction of future research. The second direction of future work is a comprehension time series functions comparative analysis in context of analysis, indexing and classification of paired TTGs. At last, the third direction of future studies is development of combined “effectiveness and efficiency” time series distance function in context of paired TTGs analysis.

As known by the author, at this time, there are no another works devoted to paired TTGs (because of its specific), particularly in context of Parkinson’s disease diagnostics. But, as soon as these research works would appear, we could propose adequate comparative analysis method based on theoretical elements represented in this paper.

References
[1] Elble R J and Koller W C 1990 Tremor (London, Baltimore)
[2] Bernstein N A 1947 On the Construction of Movements (Moscow: Medgiz) p 254
[3] Romanov S P, Pchelin M G and Iakimovskii A F 1997 Characteristics of isometrically recorded tremor in extrapyramidal disorder Russian journal of physiology 83 9 133–9
[4] Romanov S P, Aleksanyan Z A and Manoilov V V 2004 Characteristics of tremor in normal subjects and in the diagnosis and treatment of parkinsonism Neuroscience And Behavioral Physiology (Springer New York Consultants Bureau) 34 4 pp 389-98
[5] Ding H, Trajeccevsii G, Scheuermann P, Wang X and Keogh E 2008 Querying and mining of time series data: Experimental comparison of representations and distance measures PVLDB 1 pp 1542-52.
[6] Faloutsos C, Ranganathan M and Manolopoulos Y 1994 Fast Subsequence Matching in Time-Series Databases SIGMOD Conference
[7] Yi B-K and Faloutsos C 2000 Fast Time Sequence Indexing for Arbitrary Lp Norms VLDB
[8] Berndt D J and Clifford J 1994 Using dynamic time warping to find patterns in time series KDD Workshop
[9] Keogh E J and Ratanamahatana C A 2005 Exact indexing of dynamic time warping. Knowl. Inf. Syst. 7 3
[10] Chen L and Ng R T 2004 On the marriage of lp-norms and edit distance VLDB
[11] Chen L, Ozsu M T and Oria V 2005 Robust and fast similarity search for moving object trajectories SIGMOD Conference
[12] Vlachos M, Gunopoulos D and Kollios G 2002 Discovering similar multidimensional trajectories ICDE
[13] Lin J, Keogh E J, Wei L and Lonardi S 2007 Experiencing SAX: a novel symbolic representation of time series Data Min. Knowl. Discov. 15 2
[14] Aßfalg J, Kriegel H.-P., Kröger P, Kunath P, Pryakhin A and Renz M 2006 Similarity search on time series based on threshold queries EDBT
[15] Chen Y, Nascimento M A, Ooi B C and Tung A K H 2007 SpADe: On Shape-based Pattern Detection in Streaming Time Series ICDE