

An Automated Time-Shift Alignment Algorithm Based on Discrete to Continuous Approach

^{1,2}Wajih Rhalem, ¹Mourad Raji, ³Ahmed Hammouch and ¹Jamal El Mhamdi

¹LRGE Laboratory, ENSET MohammedV University, Rabat, Morocco

²National Center for Scientific and Technical Research (CNRST), Rabat, Morocco

³Director of Scientific Research and Innovation at the Ministry of Higher Education in Rabat, Morocco

Article history

Received: 9-01-2019

Revised: 7-03-2019

Accepted: 13-04-2019

Corresponding Author:

W. Rhalem

LRGE Laboratory, ENSET

MohammedV University, Rabat,

Morocco

Email: Wajih.dgapr@gmail.com

Abstract: Dealing with the variation in retention time and time series is one of the most popular challenges in the field of analysing data obtained from biological experiments. In most cases, this problem can lead to invalid conclusions. This work aims to propose a new method to cope with this problem by using the principle of the transition from discrete to continuous entities. The experimental results have shown good performance of the proposed algorithm that has been applied on different types of time series data.

Keywords: Time Series, Time Warping Similarity, Distance Signal Alignment, Discrete to Continuous

Introduction

Time series alignment has attracted a lot of interest in recent years in various fields such as sequence matching, pattern detection and sub-sequence search. It is highly important to make comparisons of chronological series which appear, very often, deformed in time by presenting a similarity of amplitude and forms with a phase shift in the time. Time shift problem is often encountered in the analysis of data obtained from biological experiments Aach and Church (2001); Clifford *et al.* (2009), online signature validation Xia *et al.* (2018), pattern recognition Keogh and Ratanamahatana (2005); Ratanamahatana and Keogh (2004); Ratanamahatana *et al.* (2010); Araújo *et al.* (2015) as well as in signal processing Barth *et al.* (2015) and the recognition of human activities Machado *et al.* (2015); Folgado *et al.* (2018). The clear majority of approaches applied in this area have been focusing on indexing based on the Euclidean metric Agrawal *et al.* (1995); Chan *et al.* (2003); Das *et al.* (1998); Debregeas and Hebrail (1998); Faloutsos *et al.* (1994); Keogh *et al.* (2001); Korn *et al.* (1997); Yi and Faloutsos (2000), which assumes that discrete signals are equidistant points in time and are also aligned in time axis. However, there is a growing awareness that Euclidean metric is a very weak measure of distance Aach and Church (2001); Bar-Joseph *et al.* (2002); Chu *et al.* (2002), since it does not allow an elastic offset of the time axis, to adapt to similar sequences that are out of phase.

To overcome the weakness of Euclidean distance, a new technique based on dynamic programming

(Dynamic Time Warping: DTW), applied in speech processing Itakura (1975); Kruskall and Liberman (1983); Rabiner *et al.* (1978) Itakura (1975); Kruskall and Liberman (1983) was introduced by Berndt and Clifford (1994) in the time series comparison field.

In order to increase the performance of this approach in the field of time series alignment, extensions were namely introduced:

- Continuous Dynamic Time Warping (CDTW) that allows the correspondence between elements not belonging to the original time vector for each series Munich and Perona (1999); Eilers (2004)
- Dynamic Time Warp Distortion Delta (DTW-D) has been proposed by Chen *et al.* (2013), in order to minimize the deformation path resulting in alignment between two signals in the continuous domain

In our case, we have chosen to study the correction of the retention time variation and the measurement of the similarity of the time series on electropherograms resulting from a high-throughput sequencer, using a new approach based on the principle of switching from discrete to continuous of entities to be treated.

An electropherogram is a graphical representation of the data received from a sequencing machine. Each line represents one of the four nucleotides and the peaks in the lines indicate the strength of the signal. In other words, each peak corresponds to a nucleotide base and the order of the peaks is the nucleotide sequence.

Figure 1 shows two raw data from two different capillaries of a sequence.

To test the effectiveness of our algorithm in the alignment of time-shifted signals, comparative tests were

performed with the DTW algorithm. Both algorithms have been implemented in the Java programming language. For the extraction of said data the Chromas software was used (Fig. 2).

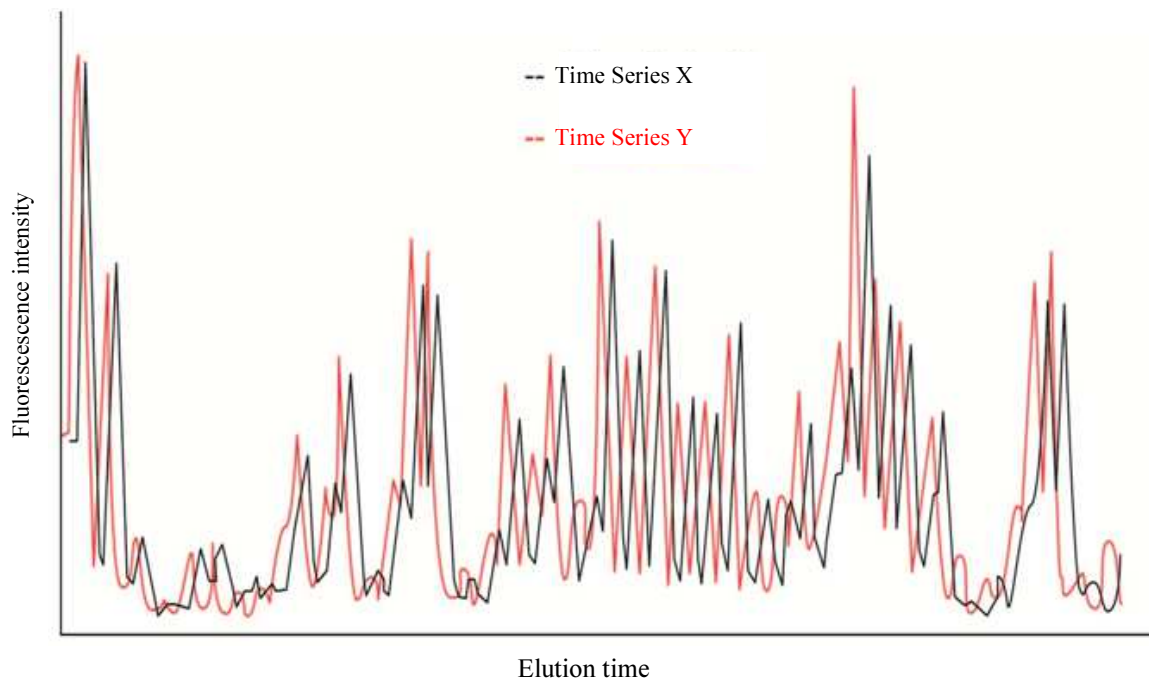


Fig. 1: Raw data obtained from two different capillaries. Time series X (black) and Time series Y (red) have the same pattern but are out of phase

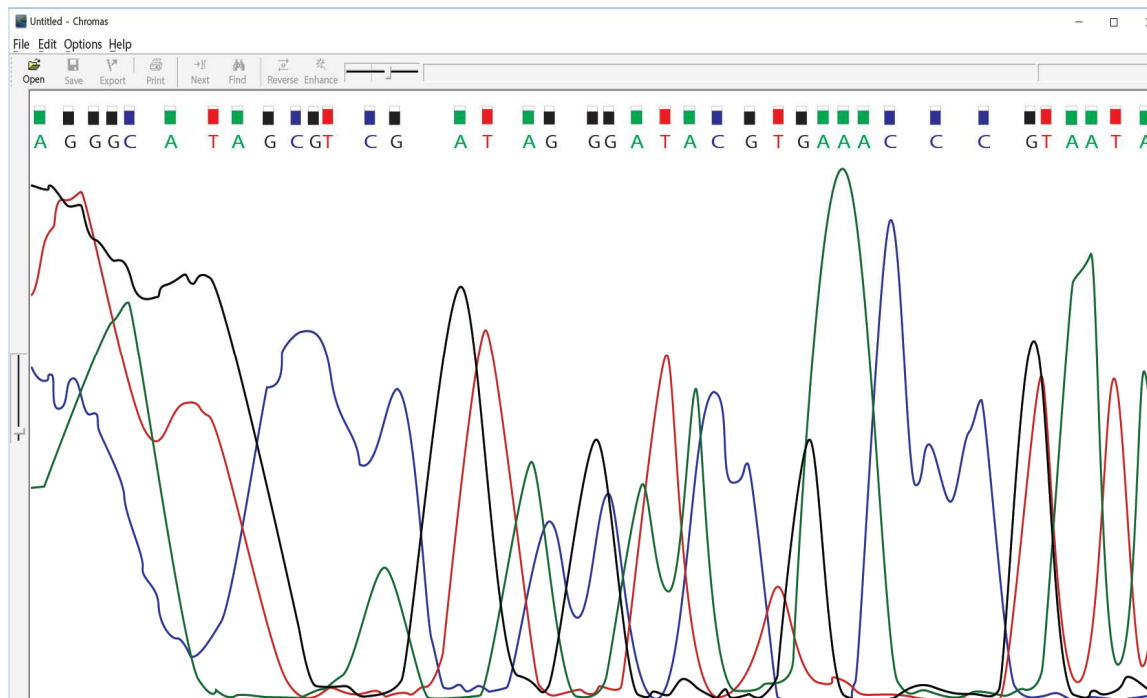


Fig. 2: Ab1 file: A biological data obtained from a high-throughput sequencer analyzed by the Chromas software

This work is organized as follows: section II presents a scientific contribution of our study, section III describes the alignment algorithms and the results obtained are presented and discussed in Section IV.

Scientific Contribution

Related Works

Given the importance of correcting the variation of retention time and time series in the field of analysis of data obtained from biological experiments and time series several studies and methods have been developed.

In addition to the conventional DTW and its previously mentioned CDTW and DTW-D extensions, a new DTW version was modified by Karabiber (2013a), which aims to improve its performance. In this direction, he also developed in the same year a new approach Karabiber (2013b) based on dynamic programming to align time series data across multiple time series. Folgado *et al.* (2018) conducted a study on performance improvement measures of time series alignment in the characterization of synthetic and real data during human movement.

In another study conducted by Robinson *et al.* (2007), a method was carried out on the basis of dynamic programming and similarity peaks to ensure the alignment of Gas Chromatography-Mass spectrometry.

Gong *et al.* (2004) conducted a study aimed at correcting the retention time of chromatographic fingerprints of herbal medicines, based on the interpolation of the cubic spline and the chemometric resolution.

In the above studies, it has been revealed that existing methods encounter the problem of the calculation cost which is generally high. Our approach is trying to achieve a better overlay with a reduced calculation cost.

Alignment System

As shown in the block diagram shown in Fig. 3, for making a final alignment decision, the raw signal at the input must undergo different stages of processing and transformation, for example pre-processing and feature extraction. This allows the processing of input data in the form of different representations with the aim to reduce the number of features and transforming the input signal into a suitable form for decision-making.

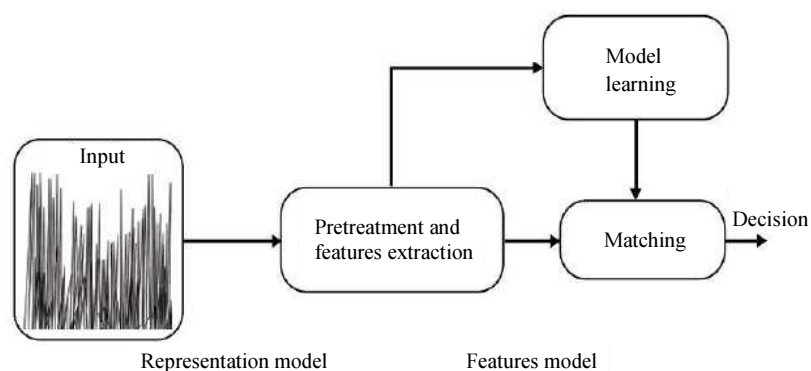
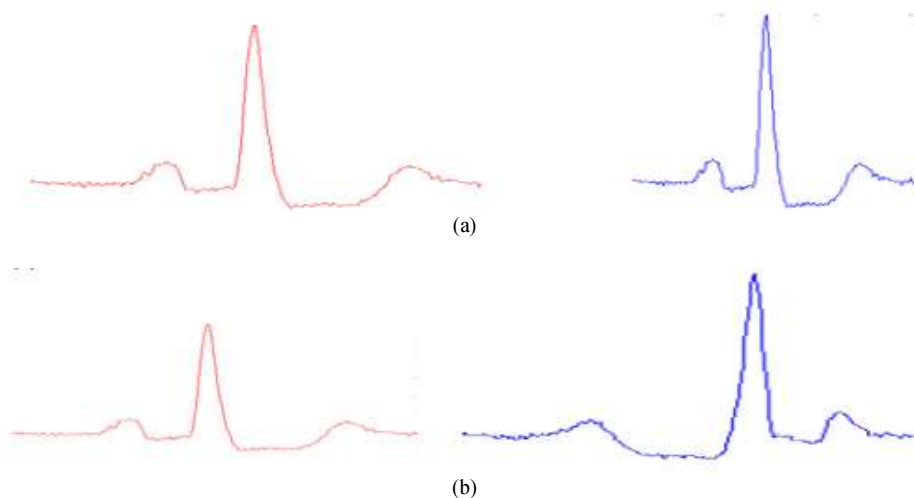


Fig. 3: Synoptic diagram of an alignment system of signals shifted in time



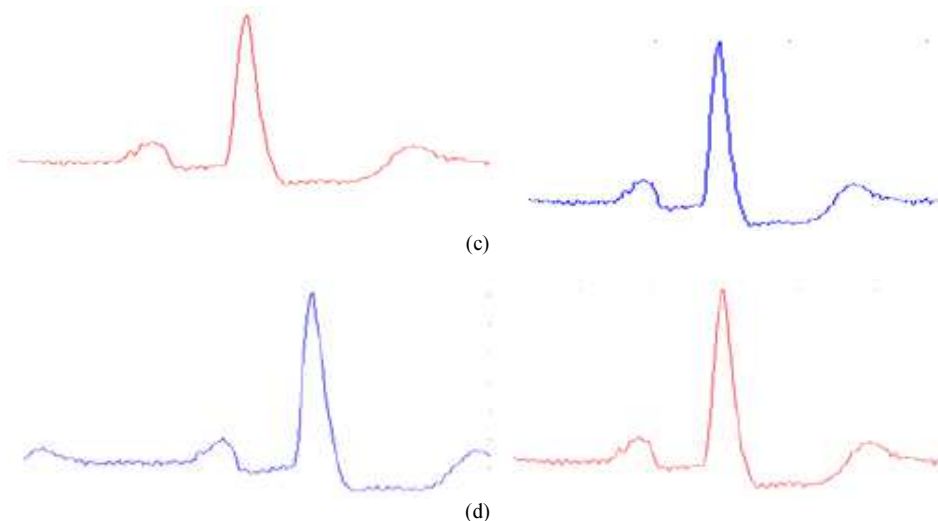


Fig. 4: Difficulties encountered in the alignment of time series. (a) difference in time scales (b) difference of amplitudes scales (c) shift on the amplitude axis (d) shift on the time axis

The superposition of the signals shifted in time, represents a major problem. For instance, we can mention: The problems of time scale differences (Fig. 4a), difference in amplitude scales (Fig. 4b), shift on the amplitude axis (Fig. 4c) and shift on the time axis (Fig. 4d). Hence the need for a good measure of similarity that allows both time-shifted signals to be compared and aligned, taking into consideration distortion problems in the data.

Signal Pre-Processing

The signals obtained from the sequencers have undergone the various stages of preprocessing and extraction of the characteristics via the tools implemented in the sequencers, which are essential for improving the performance of the alignment algorithm. These tools are summarized and presented below:

The improvement of signal comparison systems requires the implementation of the following preprocessing tools:

Smoothing

This preprocessing step is to reduce the noise of raw data using smoothing methods such as triangular method.

Signal Enhancement

In this phase, second derivative-based resolution enhancement techniques must be deployed to ensure that the total area under the curve of the derivative of a shaped signal is zero.

Adjustment of the Baseline

To write off the background signal and guarantee the normalization of the baseline an algorithm adjustment has to be deployed Vasa *et al.* (2008).

Normalization

To eliminate experimental bias, the statistical normalization of unit-to-average variance is often used. The average of the data is subtracted from each data point and these differences are then divided by the standard deviation to obtain data.

Contribution of the Discrete-To-Continuous “DTC” Algorithm

The main objective of this research is to contribute to the improvement of an alignment system based on the approach of the transition from discrete to continuous.

The DTC approach has already proved useful in several areas, such as speech recognition Maazouzi *et al.* (2015); (Robinson *et al.*, 2007), online signature matching Aqili *et al.* (2016a), fingerprint matching Aqili *et al.* (2016b) and computer geometry Aqili *et al.* (2016c).

While other algorithms try to find a good fit of certain pairs of features by comparing the composition of the two features (point by point), the DTC algorithm treats this problem as a point pattern matching problem. It brings the discrete representation of the test signal to the continuous representation (using the cubic spline interpolation) of the set of learning data by considering the problem in its entirety and not point by point. This makes it possible to compare the signals by seeking the inclusion of one form in another by searching for the best superposition.

Alignment Algorithms

After the preprocessing step of the signal (smoothing, signal enhancement and normalization), comes the step of recognition which is a main task in

shape alignment systems in terms of accuracy and execution time. The DTC has to be developed and implemented as an algorithm to find the correspondence between the entities to be processed and compare it with the tried and tested Dynamic Time Warping algorithm, commonly used in the field of time-shifted signal processing.

Dynamic Time Warping (DTW)

The DTW method Itakura (1975); Kruskall and Liberman (1983); Rabiner *et al.* (1978); Kruskall and Liberman (1983) consists in establishing a nonlinear alignment of the sequences to allow the matching of the sequences of different lengths or shifted in time.

To align two sequences $X = (x_i), i = 1: n$ and $Y = (y_j), j = 1: m$, with the DTW algorithm, we construct an accumulation matrix of distance $m*n$, each cell represents the alignment between two points x_i and y_j calculated by one of the Minkowski distances. The Euclidean distance (Equation 1) is most often used to calculate these cells:

$$d(x_i, y_j) = (x_i - y_j)^2 \quad (1)$$

Technically, the values of the sequences to be compared $X = (x_i), i = 1: n$ and $Y = (y_j), j = 1: m$ are replicated until the best match between the modified sequences is obtained.

The latter have the same element number k with $\max(n, m) \leq k \leq n+m+1$. Once the accumulation matrix of distances has been constructed, the algorithm finds the minimum alignment path through the matrix. This alignment path also called deformation path defines the correspondence between the elements of the first and second series. There are several constraints on the deformation path, such as: Boundary conditions: $w_1 = (1,1)$ and $w_K = (m, n)$, to force the breakpoints in each series to match.

Continuity: We give $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a-a' = 1$ and $b-b' = 1$, to restrict the allowed steps in the deformation path to adjacent cells.

Monotony: we give $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a-a' = 0$ and $b-b' = 0$.

To force the points of W to be spaced monotonously in time.

There are several deformation paths that fulfill these conditions but we choose the one that minimizes the deformation cost:

$$DTW(X, Y) \min \left\{ \sqrt{\sum_{k=1}^K w_k / K} \right\} \quad (2)$$

The k element of W is defined as $w_k = (i, j)_k$ such that:

$$W = w_1, w_2, \dots, w_k, \dots, w_K \quad (3)$$

Equation (4) illustrates the dynamic programming method used for the optimal path computation that represents the DTW distance. Let D be this cumulative distance up to the cell w_k :

$$D(x_i, y_j) = d(i-j) + \min \left\{ \begin{array}{l} D(i-1, j); \quad D(i, j-1); \\ D(i-1, j-1) \end{array} \right\} \quad (4)$$

As far as complexity is concerned, it has been proven that DTW's is of the order of $O(m*n)$.

Discrete To Continuous (DTC)

In comparison to most algorithms that rely on comparing the specificities (point by point) to find a match for characteristics pair, the DTC algorithm Raji and Cossé-Barbi (1999) focuses on addressing this issue as a whole and not point by point. The method used has to do with superimposing the test signal discrete representation on the reference signal continuous representation, providing two sets of points $F = \{f_i \in R^d\}_{i=1}^n$ (the model set) and $SF = \{Sf_j \in R^d\}_{j=1}^{m \leq n}$ (The data set) in the multidimensional space ND .

The main obstacle is to examine if the SF form is included in the F form. For most cases, we begin by inspecting the point-to-point correspondence between SF and F and perhaps look for an alteration that can overlay SF on F .

If the T transformation search is selected, a direct search for it, without knowing the correspondence of the SF points with those of F , can be very time consuming as a result of the large number of options to be tested (combinatory explosion).

It is to be emphasized that the presence of transformation T can definitely affirm the existence of SF in F . in this sense, the main objective of the DTC algorithm is finding transformation T so as to affirm the presence of SF in F .

The source of the combinatorial explosion appears to be from the distinct nature of the point clouds to be examined. For this to be solved, the DTC algorithm proposes to transfer one of the entities (F) representation from discrete to continuous.

In this case, with a continuous representation of F by a polynomial interpolation (SF would be retained in its discrete representation), the problem of deciding whether SF is included in F therefore becomes the search, not for T , but at the beginning of a transformation T which would bring SF back to the continuous representation of F . Therefore, the existence of T could induce the probable existence of T and will therefore confirm the inclusion of SF in F .

In short, the algorithm consists in avoiding a direct search for T but rather in searching for a transformation

T' , which would ensure the superposition of SF on the continuous representation of F .

It should be noted that the existence of such a transformation T is necessary but not sufficient to confirm the inclusion (total or partial) of SF in F . Indeed, the processing T (if it exists) has to guarantee that SF is sent back to the continuous representation of F . If for points $P_{sf_i} \in SF$ there exists a point $P_{f_j} \in F$ such that $T'(P_{sf_i}) = P_{f_j}$ then $T = T$ then SF is included in F .

The DTC algorithm is being developed to deal with arbitrary models that are defined by point cloud models in an N-dimensional (ND) space. In the case of our software, cloud models will be considered in a 2-dimensional (2D) space.

The points of F are given in the O_{xy} plane.

Let P_{xy} be the interpolation polynomial in the O_{xy} plane. For this reason, for each point $f_i = \{x_i, y_i\}$ belongs to F , we have:

$$P_{xy}(x_i) = y_i \quad (5)$$

We call this representation R .

There are different interpolation methods to represent R . To ensure that the degree of the polynomial does not depend on the size of the point cloud F , the DTC algorithm uses "cubic spline" interpolation, which is a third degree polynomial succession (piecewise interpolation). This also ensures continuity and differentiation over the entire interpolation interval.

Search for T' :

- The desired transformation T aims to bring the SF cloud back to the cubic interpolation of the F shape along the plane O_{xy} .
- The desired transformation T is expressed in homogeneous coordinates.

If we take into account (x_{SF_j}, y_{SF_j}) the coordinates of a point $j \in SF$ and (x'_{SF_j}, y'_{SF_j}) its transformation by T , the transformed points of the SF must verify R being:

$$P_{xy}(x'_{SF_j}) = y'_{SF_j} \quad (6)$$

The T' transformation parameters in the DTC have to do with a three-dimensional space, which are: Three translations (t_x, t_y, t_z), three rotations: ($\theta_x, \theta_y, \theta_z$) and three scale factors: ($\lambda_x, \lambda_y, \lambda_z$).

This study focuses on the problem of time alterations of electrophoresis (intensity as a function of elution time) where the transformation T turns out to be a single parameter function that is the translation t_x along the Ox axis:

$$x'_{SF_j} = x_{SF_j} + t_x \quad (7)$$

$$y'_{SF_j} = y_{SF_j} \quad (8)$$

The representation of R can consequently be demonstrated as follows:

$$P_{xy}(x'_{SF_j}) - y'_{SF_j} = 0 \quad (9)$$

Or:

$$\left(P_{xy}(x'_{SF_j}) - y'_{SF_j} \right)^2 = 0 \quad (10)$$

Applied to all SF points:

$$\sum_{j=1}^m \sqrt{\left(P_{xy}(x'_{SF_j}) - y'_{SF_j} \right)^2} = 0 \quad (11)$$

Consider QT as the following expression:

$$QT_{(y)} = \sum_{j=1}^m \sqrt{\left(P_{xy}(x'_{SF_j}) - y'_{SF_j} \right)^2} \quad (12)$$

Relying on this definition, the main goal now is to examine the parameters of the transformation T that minimizes the QT function.

The acquired QT function is a non-linear equation that is equally continuous and differentiable.

After the adjusting step of the SF points on the continuous representation of F (defined by T), we associate each SF point with its isomorph, which is its closest neighbor to F according to a type of distance and a pre-defined threshold (ϵ).

$P_j \in SF : T(x_{SF_j}, y_{SF_j}) - (x''_{SF_j}, y''_{SF_j})$ and if $P_i(x_{F_i}, y_{F_i})$ is the isomorph of the point P_j of SF defined by T . The root mean score (RMS) used to measure the total precision of the superposition of SF in F is:

$$RMS = \frac{1}{m} \sqrt{\sum_{j=1}^m \left(x''_{SF_j} - x_{F_i} \right)^2 - \left(y''_{SF_j} - y_{F_i} \right)^2} \quad (13)$$

At this phase, since the isomorphs of the points of SF in F are identified, it would be possible, to upgrade the superposition.

RMS is also formulated as a non-linear function. In general, the transformation T that is found provides T directly and thus the RMS solution is already found.

However, some improvements can be operated within a predetermined RMS threshold. After having determined the RMS parameter (t'_y), only SF points with a distance to their isomorphs in F , are below a predefined ϵ threshold. When all distances between SF points and

their isomorphs in F do not exceed ε , the SF is reported as included in F . If SF is stated not being included in F , the DTC algorithm can process the Larger set of Common Points (LCP) between the two signals. In fact, isomorphic pairs that fail to respect the predefined

threshold ε would be removed and a refining of the RMS value would resume for a better superposition of the other SF points onto the F points.

The processing steps of the DTC algorithm are illustrated in Fig. 5.

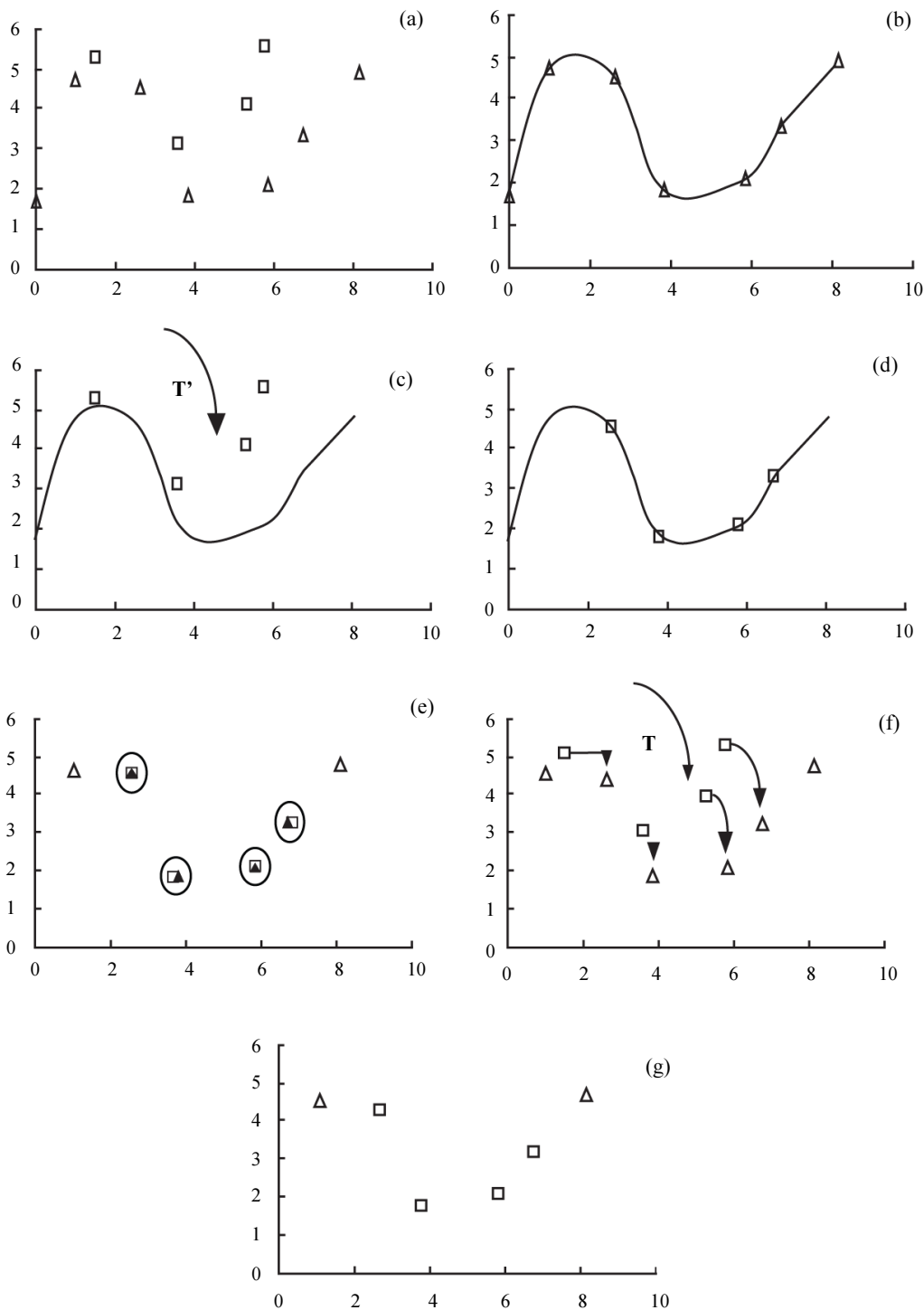


Fig. 5: The processing steps of the DTC algorithm

Algorithm Process

The following Fig. 6 gives an overview of the DTC process.

Algorithm Complexity

The study of the complexity of DTC has been done on models obtained from the generic reference database SHREC 2009 (Godil *et al.*, 2009).

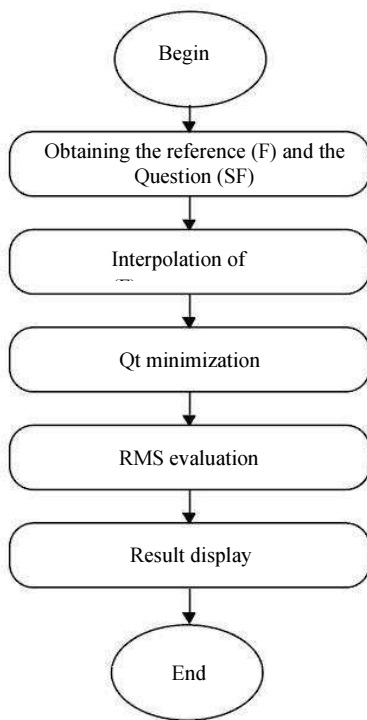


Fig. 6: DTC Process

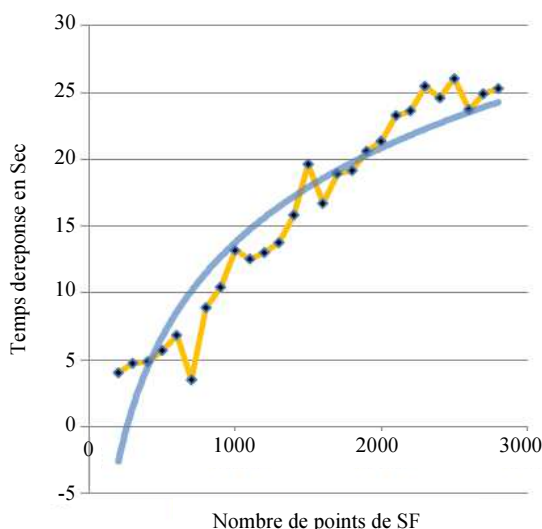


Fig. 7: Average run time for each size of SF

The experiments were conducted with the purpose of seeking matching cloud F and SF. Model F is composed of 3093 points. The initial SF size is 200 points to be increased each time by 100 points. For each SF, we apply a random transformation. This same experience is repeated 100 times. Figure 7 shows the average run time for each size of SF.

Experimental Results and Discussion

Data

To test the effectiveness of the algorithm in the processing of time-shifted signals, it was decided to generate 3 electrophoregrams of 1353 points (nucleotides) for the same DNA sample according to three different experiments using a high speed sequencer "Ion semiconductor proton sequencer".

The electrophoregrams in question were obtained from the UATRS laboratory of the National Center for Scientific and Technical Research (CNRST) in Rabat Morocco. As shown in Fig. 8, the series of signals obtained are shifted in time according to the experimental conditions.

Tools

The data obtained is in the form of an ab1. file and they were extracted by means of software of extraction and analysis of electropherograms "Chromas PRO". The algorithms described above (DTC and DTW) have been implemented using the Java programming language. The machine's features are Intel Core i7 processor at 2.40 GHz and 8 GB of RAM.

Results and Discussion

In order to test the efficiency and accuracy of our DTC algorithm, we compared it with the DTW algorithm by performing the alignment of 3 signals having a time offset problem. As illustrated in Fig. 8a, the signals have the same pattern, but the elution time and intensity are different from each other. After applying our algorithm, all time-series data are perfectly aligned (Fig. 8b).

A comparison of the two algorithms is shown in the Table 1. We observe that the DTC algorithm outperforms DTW and can correct more accurately the time shift problem.

After applying the two alignment algorithms, we notice that the execution time of the DTC algorithm (0.178s) is reduced compared to that of the DTW (0.772s), with better accuracy (87.75% for DTW against 94.66% for DTC).

Metric Similarity

The problem therefore lies in the superposition of the three signals shifted in time, if we consider that

the first signal constitutes the reference (F) and that the other two signals are the questions (SF). Our algorithm (DTC) must therefore provide us with the best alignment of these three signals. A question (SF) would be perfectly aligned with the reference (F) when the interpolation polynomial based on F would pass on the points of the transformed SF (translation t_x):

$$P_{xy}(x'_{SF_j} + tx) = y'_{SF_j} \quad (A)$$

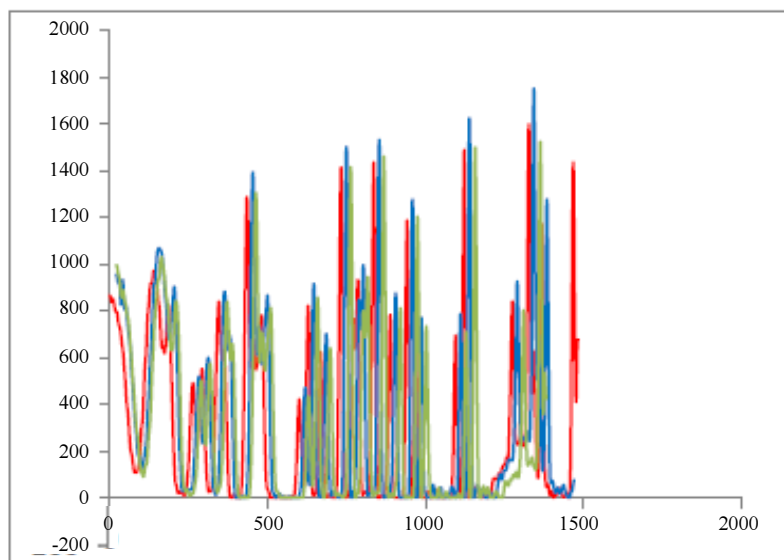
The percentage of similarity is obtained with:

$$\frac{\text{Number of } SF \text{ points checking } (A)}{\text{total size of } SF}$$

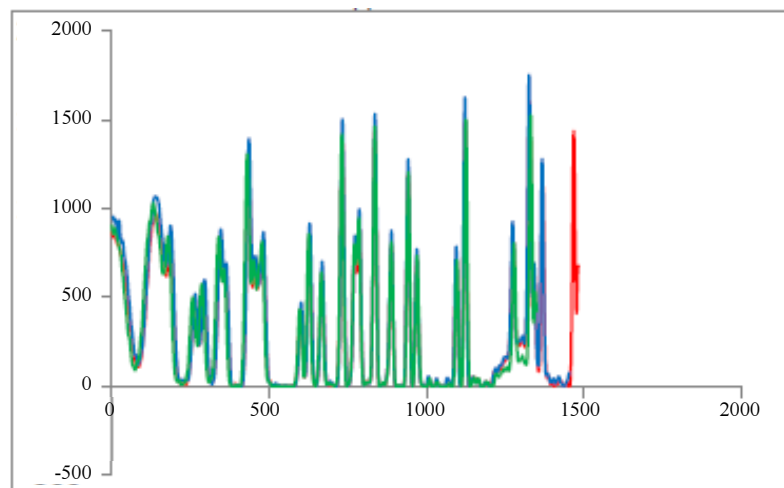
The similarity percentage of Table 1 is therefore an average of those of the two SF .

Complexity

The complexity of the space-time DTC is a logarithmic complexity with respect to the size n of the subject $O(m \log n)$ and the DTW is $O(m*n)$. Recalling that the notation O serves to demonstrate how algorithms respond to changes in input size in terms of processing time or workspace requirements. This explains the performance of our DTC algorithm.



(a)



(b)

Fig. 8: (a) 3 different signals with time shift (b) signals after alignment using DTC

Table 1: DTC and DTW comparison

Method	Complexity	Execution time	Efficiency (similarities)
DTW	$O(m*n)$	0.772s	87,75 %
DTC	$O(m \log n)$	0.178s	94.66%

Conclusion

Time-shifted signals can be corrected using the DTC approach. The main advantages of the proposed algorithm for signal alignment are speed and accuracy. Time shifted problems are solved correctly using our algorithm in a much shorter time compared to DTW.

These results encourage the development of fully automated software based on the DTC algorithm to analyze time-shifted signals.

Acknowledgement

We would like to thank the Technical Support Units for Scientific Research (UATRS) of the National Center for Scientific and Technical Research (CNRST) in Rabat, Morocco for providing the electrophoregrams used in this study.

Author's Contributions

All authors are equally contributed in this work and this paper.

Wajih Rhalem: Participated in all experiments, coordinated the data-analysis and contributed to the writing of the manuscript

Mourad Raji: Participated in all experiments, coordinated the data-analysis and contributed to the writing of the manuscript.

Ahmed Hammouch: Coordinated the mouse work, designed the research plan and organized the study.

Jamal El Mhamdi: Coordinated the mouse work, designed the research plan and organized the study.

Ethics

This article is original and contains unpublished material. The corresponding authors have read and approved the manuscript and no ethical issues involved.

References

- Aach, J. and G.M. Church, 2001. Aligning gene expression time series with time warping algorithms. *Bioinformatics*, 17: 495-508.
DOI: 10.1093/bioinformatics/17.6.495
- Agrawal, R., K.I. Lin, H.S. Sawhney and K. Shim, 1995. Fast similarity search in the presence of noise, scaling and translation in time-series databases. *Proceedings of the 21st International Conference on Very Large Databases, (VLD' 95)*, Switzerland, pp: 490-501.

- Aqili, N., A. Hammouch and M. Raji, 2016a. PPM translation, rotation and scale in d-dimensional space by the discrete to continuous approach. *Int. Rev. Comput. Softw.*, 11: 270-276.
DOI: 10.15866/irecos.v11i3.8746
- Aqili, N., A. Maazouzi, M. Raji, A. Jilbab and A. Hammouch *et al.*, 2016b. Fingerprint matching algorithm based on discrete to continuous approach. *Proceedings of the International Conference on Electrical and Information Technologies*, May 4-7, IEEE Xplore Press, Tangiers, Morocco, pp: 414-417.
DOI: 10.1109/EITech.2016.7519632
- Aqili, N., A. Maazouzi, M. Raji, A. Jilbab and S. Chaouki *et al.*, 2016c. On-line signature verification using point pattern matching algorithm. *Proceedings of the International Conference on Electrical and Information Technologies*, May 4-7, IEEE Xplore Press, Tangiers, Morocco, pp: 410-413.
DOI: 10.1109/EITech.2016.7519631.
- Araújo, T., N. Nunes, H. Gamboa and A. Fred, 2015. *Generic Biometry Algorithm Based on Signal Morphology Information: Application in the Electrocardiogram Signal*. 1st Edn., Springer, Cham, ISBN-13: 978-3-319-12610-4, pp. 301-310.
- Bar-Joseph, Z., G. Gerber, D. Gifford, T. Jaakkola and I. Simon, 2002. A new approach to analyzing gene expression time series data. *Proceedings of the 6th Annual International Conference on Research in Computational Molecular Biology*, Apr. 18-21, ACM, Washington, DC, USA, pp: 39-48.
DOI: 10.1145/565196.565202.
- Barth, J., C. Oberndorfer, C. Pasluosta, S. Schlein and H. Gassner *et al.*, 2015. Stride segmentation during free walk movements using multi-dimensional subsequence dynamic time warping on inertial sensor data. *Sensors* 15: 6419-6440.
DOI: 10.3390/s150306419
- Berndt, D. and J. Clifford, 1994. Using dynamic time warping to find patterns in time series. *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, Jul. 31-Aug. 01, AAAI Press, Seattle, WA, pp: 359-370.
- Chan, K.P., A. Fu and C. Yu, 2003. Haar wavelets for efficient similarity search of time-series: With and without time warping. *IEEE Trans. Knowl. Data Eng.* 15: 686-705.
DOI: 10.1109/TKDE.2003.1198399

- Chen, Y., B. Hu, E. Keogh and G.E. Batista, 2013. Dtw-d: Time series semi-supervised learning from a single example. Proceedings of the 19th ACM SIGKDD in Ternational Conference on Knowledge Discovery and Data Mining, Aug. 11-14, ACM, Chicago, Illinois, USA, pp: 383-391.
DOI: 10.1145/2487575.2487633
- Chu, S., E. Keogh, D. Hart and M. Pazzani, 2002. Iterative deepening dynamic time warping for time series. Proceedings of the 2nd SIAM International Conference on Data Mining. Arlington, Apr. 11-13, VA, USA. DOI:10.1137/1.9781611972726.12
- Clifford, D., G. Stone, I. Montoliu, S.F. Rezzi and P. Martin *et al.*, 2009. Alignment using variable penalty dynamic time warping. Analytical Chem., 81: 1000-1007. DOI: 10.1021/ac802041e
- Das, G., K. Lin, H. Mannila, G. Renganathan and P. Smyth, 1998. Rule discovery from time series. Proceedings of the 4th International Conference of Knowledge Discovery and Data Mining, Aug. 27-31, AAAI Press, New York, pp: 16-22.
- Debregeas, A. and G. Hebrailk, 1998. Interactive interpretation of Kohonen maps applied to curves. Proceedings of the 4th International Conference of Knowledge Discovery and Data Mining, Aug. 27-31, AAAI Press, New York, pp: 179-183. <https://dl.acm.org/citation.cfm?id=3000323>
- Eilers, P.H. 2004. Parametric time warping, Analytical Chem.. 76: 404-411. DOI: 10.1021/ac034800e
- Faloutsos, C., M. Ranganathan and Y. Manolopoulos, 1994. Fast subsequence matching in time-series databases. Proceedings of the ACM SIGMOD International Conference on Management of Data, May 24-27, Minneapolis, Minnesota, USA, pp: 419-429. DOI: 10.1145/191843.191925
- Folgado, D., M. Barandas, R. Matias, R. Martins and M. Carvalho. 2018. Time alignment measurement for time series. Patt. Recognit., 81: 268-279.
DOI: 10.1016/j.patcog.2018.04.003
- Godil, A., H. Dutagaci, C.B. Akgül, A., Axenopoulos and B. Bustos *et al.*, 2009. SHREC'09 Track: Generic shape retrieval. Proceedings of the Eurographics Workshop on 3D Object Retrieval, (DOR' 09), pp: 61-68.
- Gong, F., Y.Z. Liang, Y.S. Fung and F.T. Chau, 2004. Correction of retention time shifts for chromatographic fingerprints of herbal medicines. J. Chromatography, 1029: 173-83.
DOI: 10.1016/j.chroma.2003.12.049
- Itakura, F., 1975. Minimum prediction residual principle applied to speech recognition. IEEE Trans. Acoust. Speech Signal Process. 23: 52-72.
DOI: 10.1016/B978-0-08-051584-7.50015-2
- Karabiber, F., 2013a. A peak alignment algorithm with novel improvements in application to electropherogram analysis. J. Bioinform. Computat. Bio., 11: 1350011-1350011.
DOI: 10.1142/S021972001350011X.
- Karabiber, F., 2013b. An automated signal alignment algorithm based on dynamic time warping for capillary electrophoresis data. Turk J. Elec. Eng. Comp. Sci., 21: 851-863. DOI: 10.3906/elk-1112-71
- Keogh, E. and C.A. Ratanamahatana, 2005. Exact indexing of dynamic time warping. Knowl. Inform. Syst., 7: 358-386. DOI: 10.1007/s10115-004-0154-9
- Keogh, E., K. Chakrabarti, M. Pazzani and S. Mehrotra, 2001. Locally adaptive dimensionality reduction for indexing large time series databases. Proceedings of the ACM SIGMOD Conference on Management of Data, May 21-24, ACM, Santa Barbara, California, USA, pp: 151-162. DOI: 10.1145/375663.375680
- Korn, F., H. Jagadish and C. Faloutsos, 1997. Efficiently supporting ad hoc queries in large datasets of time sequences. Proceedings of the ACM SIGMOD International Conference on Management of Data, May 11-15, Tucson, Arizona, USA, pp: 289-300. DOI: 10.1145/253260.253332
- Kruskall, J.B. and M., Liberman, 1983. The Symmetric Time Warping Algorithm: From Continuous to Discrete. In: Time Warps, String Edits and Macromolecules: The Theory and Practice of String Comparison. Addison-Wesley.
- Maazouzi, A., N. Aqili, M. Raji and A. Hammouch, 2015. A speaker recognition system using power spectrum density and similarity measurements. Proceedings of the 3rd World Conference Complex Systems, Nov. 23-25, IEEE Xplore Press, Marrakech, Morocco, pp: 1-5.
DOI: 10.1109/ICoCS.2015.7483213
- Machado, I.P., A.L. Gomes, H. Gamboa, V. Paixo and R.M. Costa, 2015. Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization. Inform. Process. Manage., 51: 204-214. DOI: 10.1016/j.ipm.2014.07.008
- Munich, M.E. and P. Perona, 1999. Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification. Proceedings of the 7th IEEE International Conference on Computer Vision, Sept. 20-27, IEEE Xplore Press, Kerkyra, Greece, Greece, pp: 108-115.
DOI: 10.1109/ICCV.1999.791205
- Rabiner, L., A. Rosenberg and S. Levinson, 1978. Considerations in dynamic time warping algorithms for discrete word recognition. IEEE Trans. Acoust. Speech Signal Process. 26: 575-582.
DOI: 10.1109/TASSP.1978.1163164.

- Raji, M. and A. Cossé-Barbi, 1999. Shape recognition and chirality measure: Reestablishing the link between similarity and dissimilarity in discrete space. *Chemometr. Intell. Laboratory Syst.*, 47: 219-225. DOI: 10.1016/S0169-439(98)00211-1
- Ratanamahatana, C. and E. Keogh, 2004. Making time-series classification more accurate using learned constraints. *Proceedings of the 4th SIAM International Conference on Data Mining*, Apr. 22-24, Lake Buena Vista, Florida, USA, pp: 11-22. DOI:10.1137/1.9781611972740.2
- Ratanamahatana, C.A., J. Lin, D. Gunopulos, E. Keogh and M. Vlachos *et al.*, 2010. Mining Time Series Data. In: *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach (Eds.), Springer, Boston, pp: 1049-1077.
- Robinson, M.D., D.P. De Souza, W.W. Keen, E.C. Saunders and M.J. McConville *et al.*, 2007. A dynamic programming approach for the alignment of signal peaks in multiple gas chromatography-mass spectrometry experiments. *BMC Bioinform.*, 8: 419-419. DOI: 10.1186/1471-2105-8-419
- Vasa, S.M., T.N. Guex, K.A. Wilkinson, K.M. Weeks and M.C. Giddings, 2008. ShapeFinder: A software system for highthroughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*, 14: 1979-1990. DOI: 10.1261/rna.1166808
- Xia, X., X. Song, F. Luan, J. Zheng and Z. Chen *et al.*, 2018. Discriminative feature selection for on-line signature verification. *Patt. Recogn.*, 74: 422-433. DOI: 10.1016/j.patcog.2017.09.033
- Yi, B.K. and C. Faloutsos, 2000. Fast time sequence indexing for arbitrary Lp norms. *Proceedings of the 26th International Conference on very Large Databases, (VLD' 00)*. Cairo, Egypt, pp: 385-394.

Nomenclatures

DTW	<i>Dynamic Time Warping</i>
DTC	<i>Discrete To Continuous</i>
CDTW	<i>Continuous Dynamic Time Warping</i>
DTW-D	<i>Dynamic Time Warp Distortion Delta</i>
SHAPE	<i>Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension</i>
RMS	<i>Root Mean Square</i>
LCP	<i>Largest Common Point</i>