

UN MÉTODO PARA ALINEAR SERIES TEMPORALES BASADO EN CARACTERÍSTICAS DE LA ENVOLVENTE COMO PUNTO DE ANCLAJE

A METHOD TO ALIGN SEGMENTS OF TIME SERIES BASED ON ENVELOPE FEATURES AS ANCHOR POINTS

C. Jarne^{*1,2} and P. N. Alcain³

¹Universidad Nacional de Quilmes- Departamento de Ciencia y Tecnología

²CONICET

³Departamento de Física FCEN, Universidad Nacional de Buenos Aires

Recibido: 08/05/19; aceptado: 16/09/19

In the field of time series analysis, there is not a unique recipe for studying signal similarities. When having the repetition of a pattern, averaging different signals of the same nature could be complicated. Sometimes averaging is essential in the analysis of the data. Here we propose a method to align and average segments of time series with similar patterns. For this procedure, a simple implementation based on python code is provided. This analysis was inspired by the study of canary sound syllables, but it is possible to apply it in semi-periodic signals of different nature, not necessarily related to sounds.

Signal Alignment, Time series, Averaging, Python code, Open source

I Introduction

As it was defined in previous works, time series is a collection of observations made chronologically. The nature of time series includes: large in data size, high dimensionality, and the necessity to be updated continuously.¹⁻³ Many analyses in experimental research rely on the study and treatment of time series. The tasks regarding the analysis mainly can be categorized into: representation and indexing, similarity measures, segmentation, visualization, and mining.

In the context of time series data mining, a fundamental problem is how to represent the time series data.

Different mining and searching tasks can be found in the literature and they can be roughly classified into the following fields: pattern discovery and clustering, classification rule discovery and summarization.¹

Regarding the search for regularities, the partial periodic patterns are an important class in a time series analysis. The regularities can be classified into two types: (i) regular patterns: patterns exhibiting periodic behavior throughout a series with some exceptions and (ii) recurring patterns: patterns exhibiting periodic behavior only for particular time intervals within a series.⁴

Time-series mining task also requires the notion of similarity between series, based on the more intuitive notion of shape. A similarity measure should be consistent and has the following properties:

- Provide recognition of perceptually similar objects, even when they are not mathematically identical.
- Be consistent with intuition.
- Emphasize the most salient features in local and global scales.
- Be universal, in the sense that allows to identify or distinguish arbitrary objects. That means no restrictions on time series has to be assumed.
- Abstract from distortions and be invariant respect to a set of transformations.

Similarity measures can be classified into four categories. First, there are shape-based distances. They compare the overall shape of the series. Second, edit-based distances which compare two time series based on the minimum number of operations needed to transform one series into another one. The third category is feature-based distances. These include extracting the features describing aspects of the series that then are compared with any kind of distance function. The last category is Structure-based similarity: its aim is finding higher-level structures in the series to compare them on a more global scale.

Further, it is possible to subdivide this last category into two specific sub-categories: Model-based distances which work by fitting a model to the various series and then comparing the parameters of the

¹cecilia.jarne@unq.edu.ar

underlying models. On the other hand, compression-based distances which analyze how well two series can be compressed together. In this case, the similarity is reflected by higher compression ratios.

One fully extended method to compare time series is called Dynamic Time Warping (DTW). This is a well-known technique to find an optimal alignment between two given (time-dependent) sequences under certain restrictions. The sequences are warped in a nonlinear fashion to match each other.

In fields such as data mining and information retrieval, DTW has been successfully applied to automatically consider with time deformations and different speeds associated with time-dependent data.⁵

Several algorithms exist in order to implement this technique and apply it to average signals.⁶⁻⁹

Nevertheless, there is not a unique recipe that can be applied in the signal analysis. In some cases, there is the possibility to use a more heuristic approach, computationally cheap and appropriate for data with regular time patterns.

In this work, we propose a method that uses the signal envelope as a feature to align canary syllables and we provide the software implementation.

The rest of the paper is organized as follows: In Section II the method is described in general terms. Section III and IV describe the method in more detail and a link to the software implementation is provided. Finally, in Section V we present the conclusions and some further work.

II Signal alignment

We present an algorithm that allows us to segment semi-periodic patterns and estimate the average of similar patterns. To reach this aim, first it is necessary to align the individual patterns and then estimate an average amplitude and time duration. Such information is of particular interest for research an characterization of birds singing in biology.¹⁰ This method was applied to canary tonal sounds. For all Figures shown in present work, data set consist of a canary sound segment taken from a public data base provided in.¹¹

In data of this nature, different segments have differences in the time duration and amplitude, even when some features, like frequency content, remains the same.

An example to describe a canary song is shown in Figure 1. The upper panel contains a segment of the time series (the song) and the envelope obtained with the method described in.¹² The markers in purple correspond to the minimum amplitude level of the envelope. The lower panel contains the frequency content of the signal with the overlapping of the amplitude envelope in arbitrary units, with markers also at the minimum amplitude instance.

A canary can sing a repertory of different repetitions of syllables. Let us suppose that one wants to find a way to automatically measure the mean duration and amplitude of each kind of syllable. We

describe here a way to automatically cut, align and estimate these mean quantities.

To summarize in steps, the method consists of the following procedure:

1. Each sound segment is cut with a criterion that will be described in Section III.
2. To align the many instances of the same syllable, once split, we calculate the envelope.
3. An anchor point is defined with a criterion that allows aligning signals with similar patterns.
4. Finally, the signal length of each segment is re-sampled into a fixed length but preserving the individual duration of each segment in a vector.

In the following subsections, we described each step of the process.

III Segmentation of the time series (First Step)

Each sound segment is cut automatically through the use of a burly envelope obtained with a low pass filter with a cutoff frequency of the order of 20-40 Hz as shown in Figure 2. Green and purple markers indicate the maximum and the minimum value of this burly envelope. The frequency of this filter is related to the frequency of the pattern repetition. The key to this step is to find the frequency of repetition of the syllables and the minimum values to use them as an index to segment the series.

That frequency range was selected with respect to the canary syllable variations. We used the minimum of that envelope as an index to pre-cut each sound segment in individual syllables.

Envelope estimation and anchor point

This is the **Second step of the process**. For each of the segments previously obtained, the true envelope is estimated using the method detail described in.¹² During this procedure, the time duration is estimated for each segment and saved.

Now, in the **Third step**, to synchronize all syllable of the multiple segments, we used as a feature the true envelope amplitude as an anchor point. We defined as the relative maximum (or minimum) the percentage high with respect to the maximum (or minimum) amplitude in each individual syllable.

In some cases, we used the value of the relative maximum of the true envelope (the Relative High of the study shown in Figure 3), in others the relative minimum. In some cases, where the amplitude is variable or very fluctuating or even when the maximum is not well defined, we use the maximum of the low pass burly envelope as an anchor point.

For each segment, we include a margin of a certain number of frames from the anchor point to preserve the entire shape of each pulse from to the signal start to the end. In this way, we used the feature anchor point to define a starting and ending point of each segment.

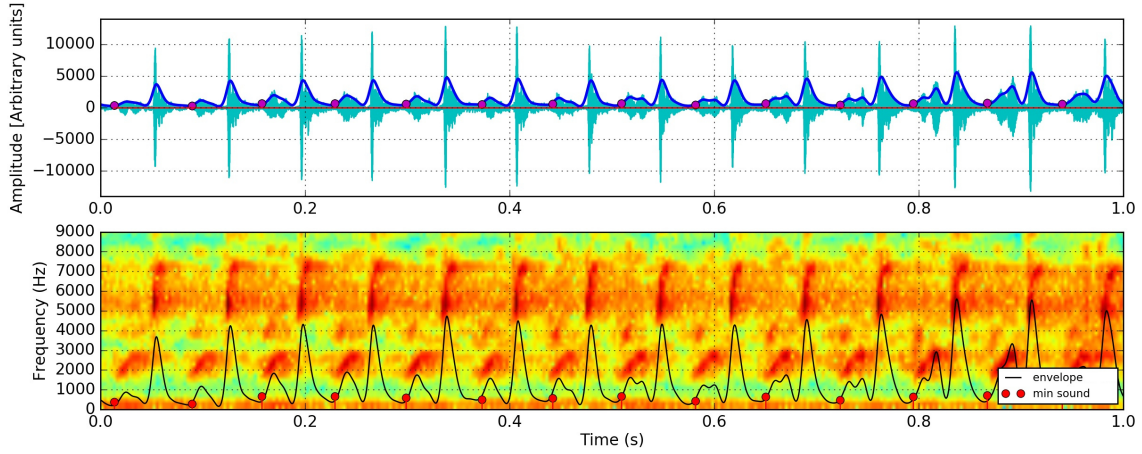


Figure 1: An example of a canary song. The upper panel shows the amplitude signal vs. time. The bottom panel shows the sound frequency content as well as the envelope in arbitrary units of amplitude. Envelope was obtained as described in.¹²

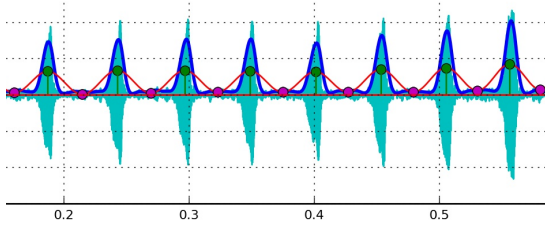


Figure 2: An example of a canary song segmentation with the burly envelope in a pink line and the signal envelope in blue. Green and purple markers indicate the maximum and the minimum value of this burly envelope.

Regarding the selection of the feature, we ask ourselves the following question: Which is the better feature to define the anchor point for a given signal segment? The answer is that it will depend on the particular characteristics of the signal. We propose the following procedure to quantify which is the better feature of a given set:

With the envelopes' amplitudes normalized to 1, we define candidates to anchor point as the earliest time t_a in which the normalized amplitude has the value a such as:

$$0 \leq a \leq 1 \quad (1)$$

We then displace all the envelopes from t to $t' = t - t_a$ in this way, all of the syllables have the same normalized amplitude at $t' = 0$

These are all the proposed alignments. For each proposed alignment, we calculate the mean squared error of the several envelopes, which is a measure of how well aligned the envelopes are. The anchor point is, then, the value that minimizes this mean squared error as shown in Figure 3.

It is also worth noting that this procedure also gives a template of the envelope of the syllables and, there-

fore, may be used to decide whether a syllable belongs to this group or not.

We applied this criteria in order to have the best possible representation or template for a given set of syllables reflected by the minimum uncertainty value that could be estimated for each considered set.

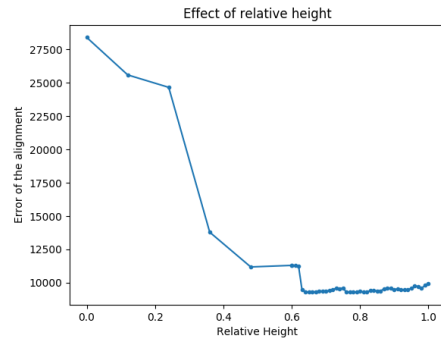


Figure 3: A study of the mean squared error vs the selected High of the anchor point.

IV Signal re-sampling and averaging. (Fourth step)

With the syllables aligned to this value, we average all the syllables. Using the anchor point, the mean value between all segment is obtained by first re-scaling a re-sampling each segment to a fixed number of points (1000 points in the case of the example in Figure 4). In this way, every segment has a fixed number of points to be aligned.

With all segments aligned, we take the average in each position. The result of this analysis is shown in Figure 5. If we want to have a mean value represented with a mean duration we can use the average of each mean time duration of the segments.¹³

In this way, when applying this procedure, we have an average syllable with an average amplitude, an average duration, and shape that could be useful for

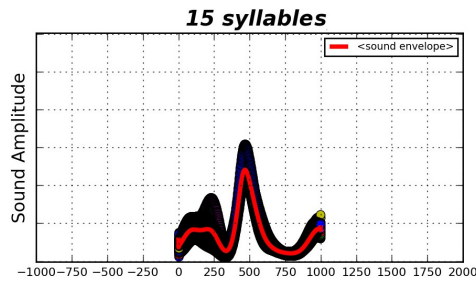


Figure 4: All segments of the song cut and overlapped. It has been re-sampled and belong to the sound time series shown in **Figure 1**. In red line the average envelope.

further segment classification as shown in Figure 5.

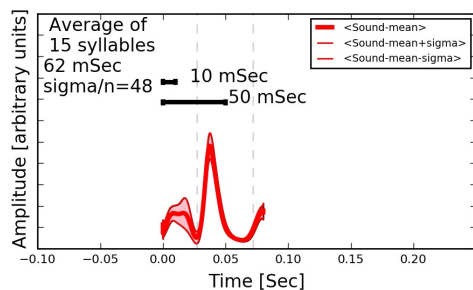


Figure 5: Mean value of the envelope corresponding to the song segments shown in **Figure 4**. Time duration corresponds to the mean value of each segment time duration.

Regarding the code implementation, the software used for this method was developed in python. Open source scientific libraries were used such as *Scipy* and *Numpy* in order to make possible to shear, modify and improve the proposed algorithm. This procedure is implemented in Python in <http://github.com/pabloalcain/syllable>

V Conclusions

In present work, we proposed a method to align and average segments or series with similar patterns. Even when the procedure is straightforward and simple, it is a useful solution. An important advantage of this method is that it can be applied in signals with different spectral content. There are no requirements on the specific domain of the data. It is appropriate for one-dimensional series of any kind. We also presented a procedure based on optimization for the anchor point selection. This procedure quantifies which anchor point, for a given, set has the minimum error. The advantage of this method presented is that it could be adapted for different signal particularities.

VI Funding

This work was developed without any additional funding other than the salary granted for FCEN and

CONICET.

VII Acknowledgments

Thank you to the reviewers who provided valuable input to each submission. I want to thank also the Stack Overflow community for the interest. Finally, I want to thank Gabriel Lio for his help.

References

1. T.c. Fu, Eng. *A review on time series data mining*, Appl. Artif. Intell. 24(1), 164 (2011). DOI : 10.1016/j.engappai.2010.09.007.
2. P. Esling, C. Agon. *Time-Series Data Mining*, ACM Comput. Surv. 45(1), 12:1 (2012). DOI : 10.1145/2379776.2379788.
3. Vetterli, J. Kovacevic, V.K. Goyal, *Foundations of Signal Processing* (Cambridge University Press, 2014)
4. R.U. Kiran, H. Shang, M. Toyoda, M. Kitsuregawa, *An Efficient Map-Reduce Framework to Mine Periodic Frequent Patterns*, in EDBT (2015) 97-108 Conf. Paper
5. F. Petitjean, A. Ketterlin, P. Gancarski, *Pattern Recognition* 44(3), 678 (2011) DOI : 10.1016/j.patcog.2010.09.013.
6. P.F. Marteau, *Times series averaging from a probabilistic interpretation of time-elastic kernel*, CoRR (2015). <http://arxiv.org/abs/1505.06897>
7. G.A. ten Holt, M.J. Reinders, E.A. Hendriks, *Multi-dimensional dynamic time warping for gesture recognition*, in Thirteenth annual conference of the Advanced School for Computing and Imaging (2007) ISBN : 9789811031564.
8. S. Salvador, P. Chan, *Toward accurate dynamic time warping in linear time and space*, Intell. Data Anal. 11(5), 561 (2007). <http://dl.acm.org/citation.cfm?id=1367985.1367993>
9. T. Giorgino, *Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package*, Journal of Statistical Software, Articles 31(7), 1 (2009). DOI : 10.18637/jss.v031.i07.
10. E.C. Tumer, M.S. Brainard, *Performance variability enables adaptive plasticity of 'crystallized' adult birdsong*, Nature (2007). DOI <https://doi.org/10.1038/nature06390>
11. Canary sound for training: https://github.com/katejarne/syllable/tree/master/test_data
12. C. Jarne, *A heuristic approach to obtain signal envelope with a simple software implementation*, ANALES AFA, [S.l.] 29 (2018). DOI : <https://doi.org/10.31527/analesafa.2018.29.2.51>
13. T. Oates, *PERUSE: An unsupervised algorithm for finding recurring patterns in time series*, in 2002 Proc. of IEEE Int. Conf. on Data Mining 2002. (2002), pp. 330-337. DOI : 10.1109/ICDM.2002.1183920