PATTERN LOCALIZATION IN TIME SERIES THROUGH SIGNAL-TO-MODEL ALIGNMENT IN LATENT SPACE

Steven Van Vaerenbergh[1], Ignacio Santamaría[1], Víctor Elvira[2] and Matteo Salvatori[3]

[1] University of Cantabria, Santander, Spain.
[2] IMT Lille Douai CRISTAL (UMR 9189), Lille, France.
[3] Tecnatom S.A., Madrid, Spain.

ABSTRACT

In this paper, we study the problem of locating a predefined sequence of patterns in a time series. In particular, the studied scenario assumes a theoretical model is available that contains the expected locations of the patterns. This problem is found in several contexts, and it is commonly solved by first synthesizing a time series from the model, and then aligning it to the true time series through dynamic time warping. We propose a technique that increases the similarity of both time series before aligning them, by mapping them into a latent correlation space. The mapping is learned from the data through a machine-learning setup. Experiments on data from non-destructive testing demonstrate that the proposed approach shows significant improvements over the state of the art.

Index Terms— Pattern localization, dynamic time warping, canonical correlation analysis, time series, alignment.

1. INTRODUCTION

The localization of patterns in a time series is a common data mining problem, with applications in a multitude of fields, including biomedicine, finance, industrial engineering and speech recognition [1, 2, 3, 4]. Contrary to the problem of detection, in which a decision is to be made about the presence or absence of a pattern, the problem of localization assumes that the pattern is present and its precise location is to be retrieved. The temporal nature of the data acquisition process complicates these tasks, as it causes the shape of the patterns of interest to suffer deformations in time known as *warps*. For pattern detection problems, many techniques exist based on aligning the query time series to a known reference pattern, commonly through dynamic time warping (DTW) [5]. Similarly, a common pattern localization technique consists in aligning the query time series to a reference time series that contains several patterns of interest [3].
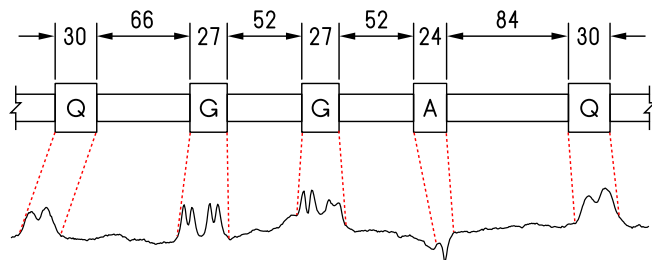
Fig. 1. Concept of the signal-to-model alignment problem. Top: A blueprint indicating the sizes (in mm) of references on a 1D body. Bottom: A time series acquired from the true physical structure. The dashed lines indicate the correct alignment between the model and the samples of the signal.

In many scenarios, however, the sequence of patterns to detect is not available as a reference time series but in the form of a blueprint that lists the theoretical locations of the expected patterns. In this case, the localization problem translates to aligning the time series to a model (the blueprint). A conceptual overview of this problem, which we refer to as "signal-to-model" (S2M) alignment, is given in Fig. 1.

The most prominent example of the S2M alignment problem in the scientific literature is "audio-to-score" alignment [6] in the context of music information retrieval. There, a musical recording is to be aligned to the corresponding score. S2M alignment is encountered in several other contexts as well, such as industrial environments, where a structure is measured in a one-dimensional fashion and the measured time series is to be aligned to a blueprint for this structure. This is the scenario illustrated in Fig. 1, and a specific application in non-destructive testing will be discussed in detail in the experiments of Section 4.

The common approach to tackling the S2M alignment problem consists in first synthesizing a time series based on the available model, and then aligning the true time series to this artificial time series through standard alignment techniques for localization [6]. The main bottleneck in this approach is the availability of an accurate synthesis technique. In some contexts, such as music, techniques for syn-

thesis are available that rely on pre-recorded waveforms and domain-specific knowledge. In other contexts, the design of an appropriate synthesis technique often represents a complex problem. An additional difficulty is due to variations in patterns that belong to the same class, illustrated by the pattern "G" in Fig. 1. Even if an appropriate synthesis technique were available, the alignment would be hindered by these differences in the true waveforms.

In this paper, we propose a localization framework based on machine learning that increases the similarity between the synthesized and the true time series by mapping both into a latent *correlation* space. The mapping is learned from a set of training signals, and by assigning a high degree of freedom to this mapping, the choice of an appropriate synthesis technique becomes less critical. We apply the proposed method to time series data acquired in non-destructive testing, obtaining a significant improvement over the state of the art.

2. SIGNAL-TO-MODEL ALIGNMENT

2.1. Problem definition

We are given a time series $X = [x_1, x_2, \ldots, x_N]$ and its corresponding theoretical model $\mathcal{M}$, consisting of a sequence of time markers $\mathcal{M} = \{(s_i, t_i, c_i)\}_{i=1,\ldots,M}$. Here, s_i and t_i mark the start and end locations of the i-th event of interest, respectively, and $c_i \in \mathcal{C}$ indicates the class of the event. The start and end locations can be expressed in any unit that represents a one-dimensional quantity, for instance millimeters or seconds. The localization problem can be solved by aligning the time series to the model $\mathcal{M}$, such that for each of the markers s_i and t_i the corresponding locations (expressed in samples) in the time series are retrieved. We will denote the alignment solution as $\mathcal{A}$.

2.2. Standard approach

The standard approach to solving S2M alignment consists in transforming it into a setting of time-series alignment. This, however, requires two time series, while the described scenario contains only one (in addition to the model). The second time series, Y, is synthesized from the information contained in the model. We briefly outline three synthesis strategies:

- *Binary synthesis* generates a time series whose values are 1 in the range $[s_i, t_i]$ of each event and 0 outside. This is a rudimentary synthesis that may serve if the true patterns resemble rectangular blocks.

- *Replication synthesis* consists in replicating a true pattern at every position indicated by the model. This type of synthesis requires the availability of a *training* signal from which patterns can be extracted.

- *Generative synthesis* encompasses advanced forms of synthesis that use domain-specific models to generate
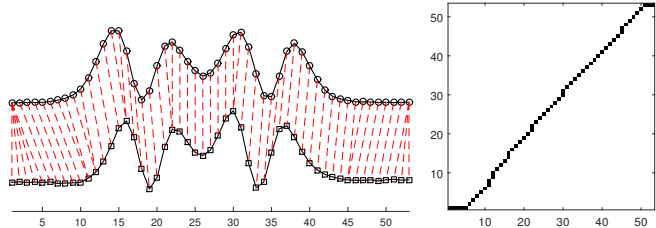

Fig. 2. Dynamic time warping. Left: Optimal alignment between two time series. Right: Corresponding warping path.

realistic time series. In music-alignment problems this type of synthesis is most often used [6].

Once the synthesized time series Y is available, it is aligned to the true time series X, typically through dynamic time warping [7, 8]. DTW is a well-known technique and the de-facto standard algorithm for aligning time series. By relying on dynamic programming, DTW allows to evaluate a combinatorially large number of warps, each of which consists of local shifts, contractions, and stretches of the signals. Figure 2 illustrates an alignment solution obtained by DTW, and its corresponding warping path. The alignment solution, denoted as $\mathcal{A}$, contains the warping path, which consists of the pairs of samples from X and Y that are aligned.

The solution $\mathcal{A}$ allows to map any location in the model $\mathcal{M}$ (across Y) to a sample index in the true time series X, and vice-versa, hence solving the localization problem.

2.3. Challenges

As mentioned in the introduction, the main bottleneck in the standard approach lies in the availability of an accurate synthesis technique. In problems with little domain-specific knowledge, for instance, only basic forms of synthesis can be applied, which may not guarantee sufficient similarity between the synthesized and the true time series. Furthermore, even if a realistic synthesis procedure were available, the patterns in the true time series may show variations that hinder the final alignment.

In order to mitigate the differences between the synthesized and the true time series, an additional transformation is sometimes performed on both. For music, a chromagram can be used, which is a condensed form of the spectral information, representing notes [6]. Nevertheless, such transformations are not guaranteed to be optimal, and in essence they require the availability of domain-specific information.

3. PROPOSED TECHNIQUE

We propose to automatically learn a transformation that maximizes the similarity between the synthesized and the true time series. This transformation can be interpreted as a mapping to a latent space where both time series are more similar. The
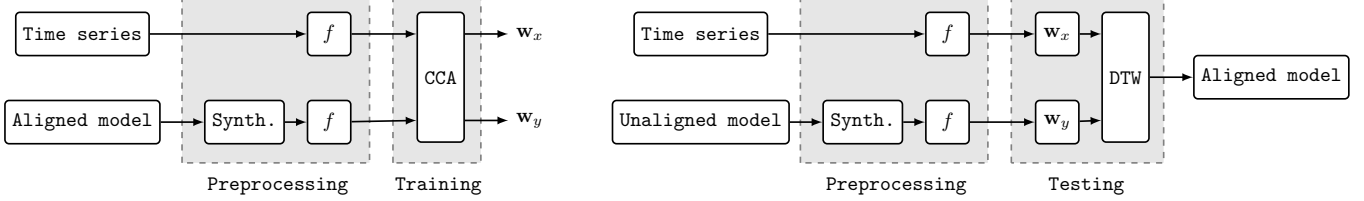
Fig. 3. Overview of the proposed machine learning framework. Left: Training setup. The block "Synth." represents the synthesis of a time series given a model $\mathcal{M}$, and f represents a predefined transformation. Right: Test setup. After preprocessing, testing consists of mapping the time series to the latent space through the projections $\mathbf{w}_x$ and $\mathbf{w}_y$, and finally DTW is performed.

goal of the mapping, therefore, is to emphasize the common components in both time series, and to suppress noise signals and artifacts present in only one of them.

The learning process is illustrated in the diagram of Fig. 3, left plot. In order to learn the mappings, a *training* time series X is required, together with a model that has been aligned to X, for instance by a human labeler. A synthetic time series Y is first obtained from the aligned model, which now represents a time series that is perfectly aligned to X.

Next, the optimal transformations need to be learned. The transformations used in this work are linear filters, which are implemented as projections of time-embedded vectors. The transformations f in the diagram, therefore, correspond to time embeddings in this case. In order to obtain the optimal mappings, the technique of canonical correlation analysis is used, which we discuss in the sequel.

3.1. Learning the latent correlation space

Given two multidimensional random variables $\mathbf{x}$ and $\mathbf{y}$, canonical correlation analysis (CCA) seeks a pair of optimal linear transformations such that the transformed variables are maximally correlated [9, 10]. In the present context, $\mathbf{x} \in \mathbb{R}^M$ and $\mathbf{y} \in \mathbb{R}^M$ represent samples from the time series that have been time-embedded with an embedding of size M. Denote by $\mathbf{w}_x$ and $\mathbf{w}_y$ the respective projection vectors, and by $x^l = \mathbf{x}^\top \mathbf{w}_x$ and $y^l = \mathbf{y}^\top \mathbf{w}_y$ the samples from the obtained transformed time series in the latent space, X^l and Y^l. If we represent the covariance matrix between $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{R}_{xy}$, and equivalently the autocovariance matrices $\mathbf{R}_{xx}$ and $\mathbf{R}_{yy}$, the function to be maximized is

$$\rho = \max_{\mathbf{w}_x, \mathbf{w}_y} \frac{\mathbf{w}_x^\top \mathbf{R}_{xy} \mathbf{w}_y}{\sqrt{\mathbf{w}_x^\top \mathbf{R}_{xx} \mathbf{w}_x \mathbf{w}_y^\top \mathbf{R}_{yy} \mathbf{w}_y}}. \qquad (1)$$

This problem can be solved as a generalized eigenvalue problem, which yields the optimal projectors $\mathbf{w}_x$ and $\mathbf{w}_y$, see [10] for additional details. Note that the training process is not limited to a single time series, as the covariance matrices can be easily constructed for an entire set of training data.

3.2. Testing on new time series

When a new time series is considered, alignment can now be performed in a straightforward fashion as illustrated by the diagram of Fig. 3, right plot. Apart from the time series, the proposed technique requires the corresponding model as an input, which is not aligned yet at this point. Similarly to the training process, the model is synthesized into a time series, and a fixed transformation f (the embedding) is applied to both time series. Then, the mapping to the latent space is performed, by projecting the true time series with $\mathbf{w}_x$ and the synthesized series with $\mathbf{w}_y$. The obtained time series in the latent space, X^l and Y^l, now show a high similarity and can be aligned by applying DTW. The alignment solution $\mathcal{A}$ is finally used to produce a model that has been aligned to the true time series (as shown in the diagram), or, equivalently, an alignment of the true time series to the original model.

The entire transformation of the time series consists in the fixed transformation f followed by the projection to the latent space. If f represents a large time embedding, $M \gg 1$, the projections possess a large number of degrees of freedom to define an optimal one-dimensional latent space. In some preliminary experiments, we have verified that this flexibility may compensate for poor forms of synthesis, such as the binary synthesis mentioned earlier. As a result, the proposed machine-learning based technique represents a framework that can be easily applied to S2M alignment problems in which domain-specific knowledge is scarce.

4. NUMERICAL EXPERIMENTS

We apply the proposed technique to time series acquired in non-destructive testing of heat generator tubes [11]. The data consists of 198 time series acquired through eddy current testing, and for each time series a corresponding blueprint is available that indicates the theoretical locations of support structures, similar to the model represented in Fig. 1. Each support structure leaves a characteristic pattern in the measured time series. Due to space restrictions, we will only consider a simplified scenario with one class of support patterns.

First, we study the basic alignment problem, in which one training time series is used to extract a pattern for replication
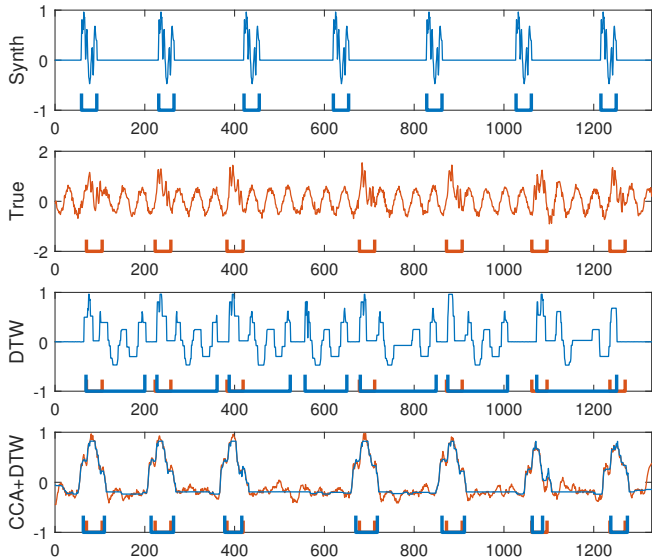
Fig. 4. Results for the first experiment. Top: Synthesized time series, with pattern starts and ends marked below. Second plot: True time series, with true starts and ends marked. Third plot: Result of aligning Synth to True using DTW; markers are shown as estimated by the alignment as well as the true markers. Bottom: Alignment of the proposed CCA+DTW technique; The latent representations for both Synth (blue) and True (red) are shown, and their markers.

synthesis, and a second time series is used for testing the localization technique. To both time series we artificially add a common background noise known as "pilgrim noise" [12], and we further add a small temporal warp. The synthesized and the test time series are shown in Fig. 4, first two plots.

As a benchmark, we apply the DTW algorithm to align both series directly. The third plot of Fig. 4 shows the synthesized time series after applying the found time warp. Clearly, the warping path does not correspond to a valid alignment. For instance, DTW determined that there is a pattern that includes sample 600 while there is none, which causes the localization of all consecutive patterns to shift. On average, the error between the true markers provided by a human labeler and markers estimated through time warping of the model amounts to 85.93 samples.

Next, we apply the proposed method with an embedding of 20 samples from the past and 20 from the future, i.e. $M = 41$. Both time series are first mapped to a latent space that is learned from the training time series, after which standard DTW is applied. The last plot of Fig. 4 shows the projections, in the latent space, of the true series and the aligned synthesized series. The estimated markers, shown below together with the true markers, have an average error of 5.86 samples.

In the second experiment, we test the alignment techniques for different degrees of pilgrim noise, ranging from 0 (no noise) to 1 (noise of the same amplitude as the maximum
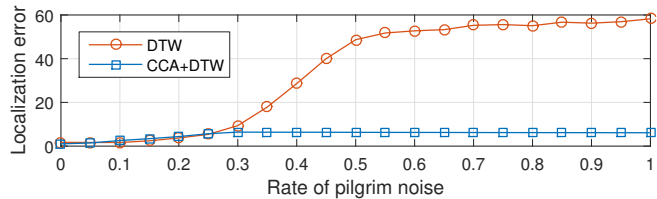

Fig. 5. Mean localization errors obtained by both methods in the second experiment, for different rates of sine noise.

peak in the signal). A single pattern was used for creating the synthesis time series, and for learning the latent space a set of 19 training time series was used. The remaining time series were used for testing. The optimal embedding for each amount of noise was determined by cross-validation on the training data.

Fig. 5 shows the average localization error for DTW and for the proposed CCA+DTW technique. For low rates of noise, both techniques perform similarly. Starting from a noise rate of 0.3, however, the benefit of operating in the learned latent space becomes clear.

5. CONCLUSIONS AND DISCUSSION

We have presented a novel technique for locating patterns in time series for which a blueprint model is available. Similar to other methods in the literature, the proposed method operates by synthesizing the model into a time series, in order to perform dynamic time warping and aligning both series. In addition, though, the proposed technique maximizes the similarity between both time series, before performing DTW, by mapping them to a latent space where the time series are maximally correlated. This optimal mapping is learned through canonical correlation analysis. Experiments with eddy current testing data show that the proposed technique is capable of correctly locating patterns in time series in challenging scenarios with high degrees of background noise.

The design of the method does not rely on any domain-specific knowledge, and is therefore expected to operate satisfactorily in a wide range of applications. In future research, we plan to apply the proposed method among others to the audio-to-score alignment problem. Due to space restrictions, we have limited the experiments in this paper to a simplified setup that highlights the strength of the proposed method. Nevertheless, preliminary experiments with more complex setups, including patterns of several different classes, multidimensional time series, and additional noise types, have been successful and will also be the topic of future research.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Tak-chung Fu, "A review on time series data mining," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164–181, 2011.

[2] Philippe Esling and Carlos Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 12, 2012.

[3] Chotirat Ann Ratanamahatana and Eamonn Keogh, "Everything you know about dynamic time warping is wrong," in *Proceedings of the 3rd Workshop on Mining Temporal and Sequential Data SIGKDD*, 2004.

[4] Zheng Liu, Jeffrey Xu Yu, Xuemin Lin, Hongjun Lu, and Wei Wang, "Locating motifs in time-series data," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2005, pp. 343–353.

[5] Donald J. Berndt and James Clifford, "Using dynamic time warping to find patterns in time series.," in *KDD workshop*. Seattle, WA, 1994, vol. 10, pp. 359–370.

[6] John Thickstun, Zaid Harchaoui, and Sham Kakade, "Learning features of music from scratch," in *International Conference on Learning Representations (ICLR)*, 2017.

[7] Lawrence R. Rabiner and Biing-Hwang Juang, "Fundamentals of speech recognition," 1993.

[8] Taras K. Vintsyuk, "Speech discrimination by dynamic programming," *Cybernetics*, vol. 4, no. 1, pp. 52–57, 1968.

[9] Harold Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[10] David R. Hardoon, Sandor Szedmak, and John Shawe-Taylor, "Canonical correlation analysis: An overview with application to learning methods," *Neural computation*, vol. 16, no. 12, pp. 2639–2664, 2004.

[11] Javier García-Martín, Jaime Gómez-Gil, and Ernesto Vázquez-Sánchez, "Non-destructive techniques based on eddy current testing," *Sensors*, vol. 11, no. 3, pp. 2525–2565, 2011.

[12] Bruno Benoist, "Expert system for eddy current signal analysis: non destructive testing of steam generator tubings," Tech. Rep. CEA-R-5553, CEA Centre d'Etudes Nucleaires de Saclay, 1991.