MULTI-LAG PHASE SPACE REPRESENTATIONS FOR TRANSIENT SIGNALS CHARACTERIZATION

Cindy Bernard, Teodor Petrut, Gabriel Vasile and Cornel Ioana

Gipsa-lab, Universite de Grenoble, Grenoble-INP,
11 rue des mathematiques,
Domaine Universitaire, BP 46,
Saint Martin d'Heres 38402 FRANCE

ABSTRACT

Transient signals are very difficult to characterize due to their short duration and their wide frequency content. Various methods such as spectrogram and wavelet decomposition have already been extensively used in the literature to detect them, but show limits when it comes to near similar transients discrimination. In this paper, we propose the multi-lag phase space analysis as a way to characterize them. This data-driven method enables the comparison between features extracted from two different signals. In an example, we compare the multi-lag phase space representations of three similar transients and show that common features can be found to discriminate them. Finally the results are compared with a wavelet decomposition.

***Index Terms*—** Phase space representation, Transients, Recurrence

1. INTRODUCTION

Transient signals are characterized by very short durations over the observation and they often characterize fast changes of the analyzed phenomena. Such signals are encountered in the analysis of various physical environments using ultrasonic or electric signals. Generally, the analysis of transient signals addresses the issues of detection and characterization. While detection of transient signals is the object of major contributions in signal processing [1] [2] [3] [4] [5], characterization of transients, e.g. for classification purposes, is a topic of growing importance in many fields such as biomedical, ultrasonic, seismic, etc [6].

In this context, this paper proposes a new representation space of transient phenomena, defined by the multi-lag phase space analysis. We show that by exploiting the multi-lag diversity of the phase diagram [7], it is possible to define a discriminant description of the transients, useful for identification or classification purposes. The actual multi-lag definition of space diagrams relies on the analysis of the signal at different scales, but this analysis is data-driven and no scale function definition is necessary. That is, the proposed concept might be interpreted as a non-parametric multi-resolution analysis. The results on simulated near similar transients prove the efficiency of the proposed concept in terms of transient characterization.

Section 2 recalls the concept of phase space recurrence for a given time delay. Then, Section 3 introduces the concept of multi-lag phase space representation and its use for the characterization of the transient signals. The example illustrated in the Section 4 shows the interest of the proposed concept. Finally, Section 5 ends this paper with some conclusions and perspectives.

2. PHASE SPACE RECURRENCE

Phase space recurrence comes from dynamical systems theory [7], [8] and takes advantage of a system's ability to return to a previously visited state. The concept enables to find out recurrence patterns that could happen in a time series. It is possible to distinguish between two different signals with similar spectra and histograms thanks to the study of those recurrence patterns. To do so, the analyzed signal $s(t)$ is first turned into a phase space trajectory $\overrightarrow{v}[s,\tau]$ by regrouping samples as vectors: this is the time-delay embedding process. The phase space trajectory at instant t corresponds to a set of m signal values chosen at different time instant:

$$\overrightarrow{v_t}[s,\tau] = \left[s_t, s_{t+\tau}, s_{t+2\tau}, ..., s_{t+(m-1)\tau}\right] \tag{1}$$

with m the embedding dimension (i.e. the dimension of the phase space), τ the time delay (or lag when dealing with samples) between successive components, and s_t the value of $s(t)$ at instant t. Once the trajectory is drawn, a recurrence is enlightened each time the trajectory intersects itself.

While it is possible to perform the phase space representation with whatever wanted embedding dimension, we restrain ourselves to the study of 2-dimension phase space vectors in order to facilitate the visualization and interpration of our results, i.e. the study will be performed with:

$$\overrightarrow{v_t}[s,\tau] = (s_t, s_{t+\tau}) \tag{2}$$

The following notation is also set down :

$$\left\{ \begin{array}{l} y_i = s\,(i+\tau) \\ x_i = s\,(i) \end{array} \right. \tag{3}$$

where i represents the i-th sample of the time series $s\,(n)$ and τ the lag.

Let us now consider three signals $s_1\,(t)$, $s_2\,(t)$ and $s_3\,(t)$ related such that:

$$s_2\,(t) = s_1\,(t+\delta) \tag{4}$$

$$s_3\,(t) = \alpha s_1\,(t) \tag{5}$$

where δ is a time delay and α a constant that modifies $s_1\,(t)$'s amplitude.

By performing the time-delay embedding process, the following properties can be noted:

$$\overrightarrow{v_t}\,[s_2, \tau] = \overrightarrow{v_{t+\delta}}\,[s_1, \tau] \tag{6}$$

$$\overrightarrow{v_t}\,[s_3, \tau] = \alpha \overrightarrow{v_t}\,[s_1, \tau] \tag{7}$$

The time-delay embedding process offers interesting properties under scaling transform [9] that will be usefull to highlight similarities between signals such as time translation and amplitudes changes.

For a given time delay, a unique phase space representation is obtained that would be different if the time delay differs [10], and it is not yet possible to know which τ would be the most appropriate to study a given signal. This is why an automatic method needs to be elaborated to select this parameter. Next Section is dedicated to the study of a complete set of phase space representations.

3. MULTI-LAG PHASE SPACE REPRESENTATION

Multi-lag phase space representation (MLPS) is the combination of different phase space trajectories obtained for different lag values:

$$MLPS\,[s, \Gamma] = \{\overrightarrow{v_t}\,[s, \tau_k]\}_{k \in [1,\ldots,N]} \tag{8}$$

where Γ, the set of lags used to perform the MLPS, is defined as follows:

$$\Gamma = \{\tau_k\}_{k \in [1,\ldots,N]} \tag{9}$$

with N the cardinal of Γ.

The MLPS can be performed with any time-embedding dimension wanted, although, our study is performed considering a two dimension time-delay embedding process (i.e. $m = 2$) in order to simplify the results's interpretation. Studied signals are defined as follows:

$$s\,(t) = \left\{ \begin{array}{lll} cos\,(2\pi f t)\,h\,(t) & \text{for} & t \in \Delta \\ 0 & \text{otherwise} & \end{array} \right. \tag{10}$$

with $h\,(t)$ an arbitrary short-time window function (in the present case, a Hamming window is considered but this study is general for any other type of window), Δ a time interval and f a given frequency.

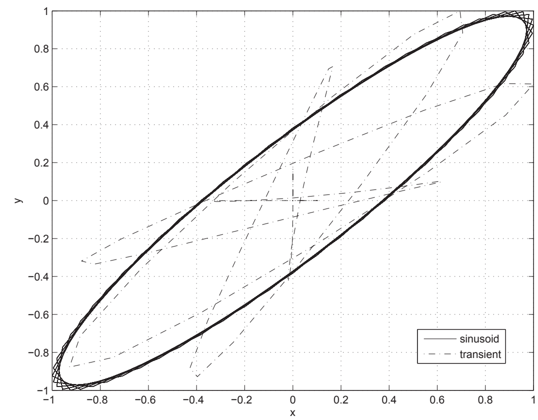

Fig. 1. This Figure displays two phase space trajectories using a given lag $\tau = 12$. The continous and dash-dotted lines respectively represent a sinusoid and a transient (defined as in Equation 10). It is obvious that even if the two signals have a similarity in common, their trajectories are completely different and allow an easy discrimination.

It is obvious that two different signals have different phase space diagrams. Figure 1 presents the diagrams of two signals having a similarity: the first one is a transient defined as in Equation 10 and the other one is a sinusoid built with the same frequency. By using any lag, the phase space diagram enables one to discriminate the two signals. However those two signals have a similarity in common, is it possible to find a way to enlight this similarity?

For a given lag τ_k, a single phase space trajectory $\overrightarrow{v}\,[s, \tau_k]$ is obtained which can be analyzed as a scatterplot C_k. The set of data is centered around the zero value and presents a trend which can be modeled by a polynomial. According to τ_k, it tends towards rotating around the central value of the distribution. The scatterplot's trend is then modeled as a third degree polynomial in order to quantify the rotation and other specificities of the trend:

$$\widehat{y}\,(t) = \widehat{a}x^3\,(t) + \widehat{b}x^2\,(t) + \widehat{c}x\,(t) + \widehat{d} \tag{11}$$

A least squares fitting estimation is chosen to model it. This is performed by minimizing the following sum:

$$\underset{\widehat{a},\widehat{b},\widehat{c},\widehat{d}}{Argmin} \sum_{i=1}^{N} (y_i - \widehat{y_i})^2 \tag{12}$$

At this point of the study, C_k's trend has been modeled by 4 parameters $\widehat{a_k}$, $\widehat{b_k}$, $\widehat{c_k}$ and $\widehat{d_k}$. Because the studied signals have a zero mean, the parameter $\widehat{d_k}$ is also equal to zero. Moreover, scatterplots present a symmetry point which cancels the parameter $\widehat{b_k}$. As a matter of consequence, C_k is now only modeled by the two remaining parameters.

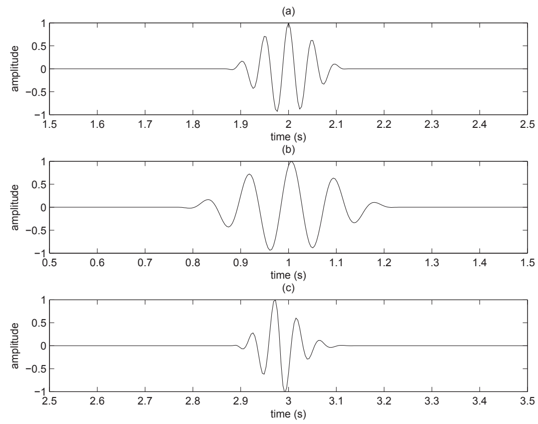

Fig. 2. This Figure represents the three studied signals. (a) represents $s_1(t)$, (b) its dilated version $s_2(t)$ and (c) is $s_3(t)$ created after a high pass filtering and a non-uniform amplitude modification of $s_1(t)$

Seeing that a scatterplot C_k corresponds to a given lag τ_k, the evolution of the parameters $\widehat{a_k}$ and $\widehat{c_k}$ can be observed according to the evolution of the lag values. The multi-lag trajectory, denoted as $MLT[s, \Gamma]$, is then defined as follows:

$$MLT[s,\Gamma] = \left[\widehat{A}, \widehat{C}, \Gamma\right] \tag{13}$$

with:

$$\begin{cases} \widehat{A} = \{\widehat{a_k}\}_{k \in [1,\ldots,N]} \\ \widehat{C} = \{\widehat{c_k}\}_{k \in [1,\ldots,N]} \end{cases} \tag{14}$$

This trajectory MLT enables to discriminate two transients as it can be studied on three different plans: $(\widehat{a}, \widehat{c})$, $(\widehat{a}, \Gamma)$ and $(\widehat{c}, \Gamma)$ plans. Next Section is dedicated to the study of an example.

4. TRANSIENTS CHARACTERIZATION

Let consider three signals :

- $s_1(t)$ is defined as in Equation 10
- $s_2(t)$ is a dilated version of $s_1(t)$ with a dilatation coefficient $\alpha = 1.8$
- $s_3(t)$ is the result of a non-uniform amplitude modification and high pass filtering successively applied on $s_1(t)$.

The temporal data is presented in Figure 2. By observing the signals, it can be noticed that while $s_1(t)$ and $s_2(t)$ have the same amplitude, their frequencies are proportional. In the same way, $s_1(t)$ and $s_3(t)$ share close frequency contents but $s_3(t)$'s amplitude has been distorted. Signals $s_2(t)$ and $s_3(t)$ have both features in common with $s_1(t)$. The object of the present study is to show that it is possible to extract those features and be able to compare them to characterize transients.

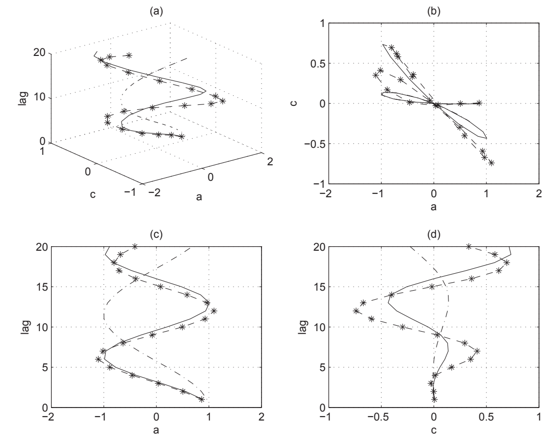

Fig. 3. This Figure represents the three MLT represented in different systems of observation. The black continuous line stands for $MLT[s_1]$, the black dash-dotted line $MLT[s_2]$ and $MLT[s_3]$ is represented by the dashed line. One can observe the scatterplots through 4 different perspectives: in three dimension in (a), in the $(\widehat{a}, \widehat{c})$ plan (b), in the $(\widehat{a}, \Gamma)$ plan (c) and in the $(\widehat{c}, \Gamma)$ plan (d).

The multi-lag phase space recurrence is performed using lags varying from 1 to 20 samples (in this example the interval Δ covers about 65 samples, meaning that performing the samples embedding process with more than 20 samples would not be significant) for the three signals. The sets of parameters $\widehat{A_i}$ and $\widehat{C_i}$ $(i \in 1, \ldots, 3)$ are then extracted from the three trajectories modeling. The 3-dimensional MLT are presented in Figure 3.

As a matter of fact, the three MLT are different but in Figure 3 (b) which shows the $(\widehat{a}, \widehat{c})$ plan, it can be seen that $s_1(t)$ and $s_2(t)$'s MLT somehow overlap which indicates that those trajectories have similarities: for lags that are different, their scatterplot's trends can be modeled by the same third degree polynomial. This is explained by the fact that one of the signal is the dilated version of the other.

In order to compare the similarity between those scatterplots, the loops that overlap are modeled by an ellipse(Figure 4). To do so, the main loops coordinates are first extracted (the longest one in term of samples) in the $(\widehat{a}, \widehat{c})$ plan and are then modeled. They are now resumed by only four parameters: their center coordinates and two integers: d_a and d_b respectively representing the semi-major and semi-minor axis.

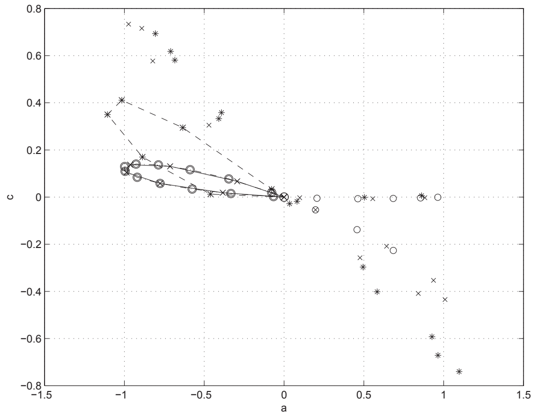

Fig. 4. This Figure represents the three MLT sketched in the $(\widehat{a}, \widehat{c})$ plan. The crosses and the continuous line are for $MLT\,[s_1]$, the circles and dash-dotted line for $MLT\,[s_2]$ and the asterisks and dashed line for $MLT\,[s_3]$. The lines correspond the selected loops which enable the comparison between signals.

An ellipse centered around 0 which major axis is parallel to the x axis can be modeled as follows:

$$\begin{aligned} F\,(x,y) &= 0 \\ \textit{with} \quad F\,(x,y) &= \Gamma x^2 + \Lambda y^2 - 1 \end{aligned} \tag{15}$$

with $\Gamma = \frac{1}{d_a^2}$, $\Lambda = \frac{1}{d_b^2}$ and $d_a > d_b > 0$.

To estimate d_a and d_b, a least square estimation is performed by minimizing the following sum:

$$\underset{\Gamma,\Lambda}{Argmin} \sum_{i=1}^{M} F^2\,(x,y) \tag{16}$$

with M the cardinal of samples used to perform the least square fitting (number of samples selected in the loop).

In the present study, the loops are not zero centered and their major axis is not parallel to the $\widehat{a}$ axis. The data is first centered and rotated, then a least square estimation is performed for the three signals. The results are shown in Figure 5. The ellipse models obtained for $s_1\,(t)$ and $s_2\,(t)$ are almost identical: their centers are really close, as their semi-major and semi-minor axis and their major axis inclination. On the contrary, the ellipse model obtained for $s_3\,(t)$ is very different. Experiment results are described in Figure 6. They do not have close centers, the representative axis are not of the same length and their major axis are different. As a way of conclusion, method has been able to enlight the fact that there exists a close relationship between $s_1\,(t)$ and $s_2\,(t)$.

In order to provide a comparison with a method of reference, a wavelet decomposition of the three signals is performed using 'Morlet' wavelet. The results are presented in Figure 7. As there is no doubt that $s_2\,(t)$ is different from

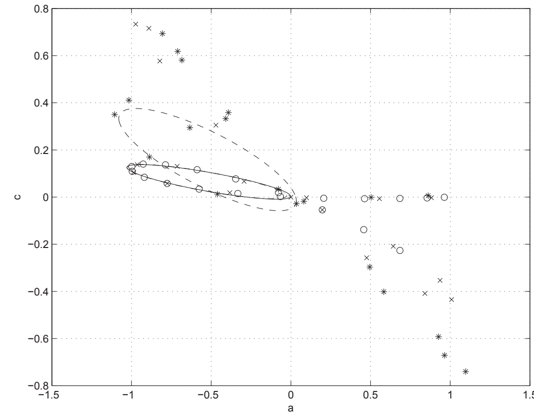

Fig. 5. This Figure represents the three MLT sketched in the $(\widehat{a}, \widehat{c})$ plan. The crosses are for $MLT\,[s_1]$, the circlesfor $MLT\,[s_2]$ and the asterisks for $MLT\,[s_3]$. The ellipses represent the modeling of the selected loops.

Signal	Distance center	d_a diff	d_b diff	angle
s_2	1.7e-4	2.38%	1.01%	0.68%
s_3	9e-3	13.18%	202.42%	167.25%

Fig. 6. This table presents the results obtained for $s_2\,(t)$ and $s_3\,(t)$ in regards of those obtained for $s_1\,(t)$. The first column displays the euclidian distance between the scatterplots centers. The second and third columns display the percentage difference between the semi-major and semi-minor axis obtained with the least square estimation. The fourth column represents the percentage difference between the inclinations of the ellipses major axis.

the other two, the results for $s_1\,(t)$ and $s_3\,(t)$ are really similar and it is difficult to say at first look that they are very different. The wavelet decomposition does not allow to discriminate transients as well as the multi-lag representation.

5. PERSPECTIVES AND CONCLUSION

The problem of transient detection has already been well covered by the literature and in particular by using phase space diagrams. Indeed those diagrams enable the extraction of pattern reccurences which can be assimilated to transients occuring in time serie (for example, electric arcing). In relation to previous work, it can be noted that Birleanu et al. have developped VeSP (vector samples processing) based tools which revealed to be very useful for transient detection, noise reduction and fundamental frequency estimation [10]. They also showed that it was possible to analyze attenuations and dilatations happening in signals thanks to recurrence plot analysis.

This paper extended the phase space representation con-

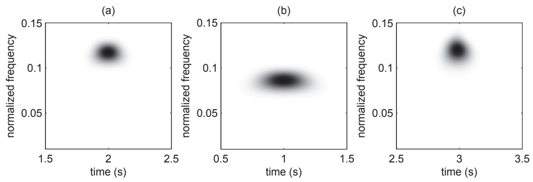

Fig. 7. This Figure represents the wavelet transform of $s_1(t)$ (a), $s_2(t)$ (b) and $s_3(t)$ (c). While the discriminate between $s_1(t)$ and $s_2(t)$ is quite obvious, it is not that simple to differentiate $s_3(t)$ from $s_1(t)$ which share a close frequency content.

cept by combining representations obtained with different lag values and proposed a method to extract features from the signals enabling the discrimination or the connection of two near similar transients. Future works will propose a method to find out more features to extract and also works on real world signals.

6. REFERENCES

[1] P. Abry and P. Flandrin, "Multiresolution transient detection," in *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis*. pp. 225–228, IEEE.

[2] Leon Cohen, *Time-frequency analysis*, Prentice Hall, 1995.

[3] S. G. Mallat, *A wavelet tour of signal processing*, Academic Press, San Diego, 2nd edition, 1999.

[4] Philippe Ravier and Pierre-Olivier Amblard, "Wavelet packets and de-noising based on higher-order-statistics for transient detection," *Signal Processing*, vol. 81, no. 9, pp. 1909–1926, Sept. 2001.

[5] Bertrand Gottin, Cornel Ioana, Jocelyn Chanussot, Guy D'Urso, and Thierry Espilit, "Detection and Localization of Transient Sources: Comparative Study of Complex-Lag Distribution Concept Versus Wavelets and Spectrogram-Based Methods," *EURASIP Journal on Advances in Signal Processing*, vol. Volume 200, Dec. 2009.

[6] B. Boashash and P. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 11, pp. 1829–1841, 1990.

[7] J.-P Eckmann, S. Oliffson Kamphorst, and D Ruelle, "Recurrence Plots of Dynamical Systems," *Europhysics Letters (EPL)*, vol. 4, no. 9, pp. 973–977, Nov. 1987.

[8] N. Packard, J. Crutchfield, J. Farmer, and R. Shaw, "Geometry from a Time Series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, Sept. 1980.

[9] Florin-Marian Birleanu, Cornel Ioana, Cedric Gervaise, Jocelyn Chanussot, Alexandru Serbanescu, and Gheorghe Serban, "On the recurrence plot analysis method behaviour under scaling transform," in *2011 IEEE Workshop on Statistical Signal Processing (SSP2011)*, June 2011, pp. 789–792.

[10] Florin-Marian Birleanu, Ion Candel, Cornel Ioana, Cedric Gervaise, Alexandru Serbanescu, and Gheorghe Serban, "A Vector Approach to Transient Signal Processing," in *11th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2012)*, July 2012, pp. 1174–1179.