ANALYSIS OF UNDERWATER SIGNALS WITH NONLINEAR TIME-FREQUENCY STRUCTURES USING WARPING-BASED COMPRESSIVE SENSING ALGORITHM

Cindy Bernard, Cornel Ioana

University of Grenobles-Alpes, GIPSA-lab
11 rue des mathematiques,
Saint Martin d'Heres, 38402 FRANCE

Irena Orovic, Srdjan Stankovic

University of Montenegro,
Faculty of electrical engineering
Podgorica, 81000 MONTENEGRO

ABSTRACT

Natural signals are often characterized by nonlinear time-frequency structures and more especially in underwater context. Underwater mammal vocalizations or dispersive phenomena are just some examples of contexts where nonlinear time-frequency structures of signal's components exist. Their analysis is of great importance for detection and classification purposes but also for phenomenon characterization.

In this work, starting from the concept of warping-based time-frequency analysis, we propose a new analysis method that combines the properties of the waping transform with the concept of compressive sensing. It provides a more accurate characterization of nonlinear time-frequency structures in terms of the estimation of their parameters. Results provided for simulated data prove the interst of this new approach with respect to the spectrogram-based method.

Index Terms— time-frequency analysis, warping transform, compressive sensing, non-linear time-frequency structures

1. INTRODUCTION

Analysis of acoustic underwater signals is an important task in many applications such as detection of underwater mammals, classification, localization, etc. The problem of analysis in passive context is a complex one while there is a large diversity of signals types more or less affected by the propagation constraints. In this paper, we consider the complex case of natural signals composed by many nonlinear time-frequency structures and that propagate in multi-path configurations. At the receiver level, the analysis of such signals is very complicated as the nonlinear time-frequency components can superpose in the time-frequency plan, creating cross-terms that could be interpreted as new (but false!) time-frequency components.

The algorithm proposed in this paper consists in developing a combined methodology of the warping operators, that stationnarize any nonlinear frequency modulation (in our works, we used mostly power class of linear modulations, but the work can be generalized to any other modulation type), and the compressive sensing-based estimation of the spectral lines that appear in the time-frequency domain after the warping operation. In Section 2, we present the general method that first consists on a warping transform that turns the nonlinear component of interest into a pure sinusoid that is sparse in the Fourier domain. This operation is then followed by a cleaning of the spectrogram using the L-statistic approach and a CS-based reconstruction algorithm that reconstructs the sinusoid. An inverse warping is finally used to express the signal in the original time domain. Afterwards, Section 3 proposes an illustration of the method with a numerical example and finally, Section 4 provides some conclusions and perspectives of work.

2. METHOD PRESENTATION

This article focuses on the recovery of signals having nonstationary time-frequency content defined as follows:

$$s(t) = \sum_{i=1}^{N} s_i(t) = \sum_{i=1}^{N} A_i e^{j\psi_i(t)} \qquad (1)$$

where N is the number of components, A_i their amplitudes and $\psi_i(t)$ their instantaneous frequency laws that are nonlinear and that can be modeled by logarithm phase signal such as:

$$\psi_i(t) == 2\pi f_{0i} t + c_i ln(t)\, ; \; t \in [t_{0i}, t_{0i} + D_i] \qquad (2)$$

with f_{0i} the central frequency of the modulation, c_i the logarithm modulation rate, t_{0i} the time instant when the modulation happens and D_i its duration.

2.1. Warping transformation

2.1.1. General formulation of a warping transformation

Generally, a warping transformation enables to express a given signal $s(t)$ into another time domain (the warped domain) that simplifies the processing of the data.

978-0-933957-43-5 ©2015 MTS

Let consider a signal $s(t) \in \mathcal{L}(\mathbb{R})$ and $\mathcal{W}$ the warping operator defined as follows:

$$\{\mathcal{W}, \quad w(t) \in \mathcal{C}^1, \quad w'(t) > 0 : s(t) \rightarrow \mathcal{W}s(t)\} \tag{3}$$

where $\mathcal{C}^1$ corresponds to the class of derivable functions. The modifications created on the studied signal are expressed through this formulation:

$$\mathcal{W}s(t) = \left|\frac{dw(t)}{dt}\right|^{1/2} s(w(t)) \tag{4}$$

where $w(t)$ is a smooth one-to-one function called the warping function. The first term $\left|\frac{dw(t)}{dt}\right|^{1/2}$ corresponds to the envelop of the warped signal that enables to conserve the energy of the original signal. It is also possible to define a non-unitary version of the warping operator if we do not want to deal with the envelop term. It is defined as follows:

$$\mathcal{W}s(t) = s(w(t)) \tag{5}$$

In most cases, warping functions are chosen in order to stationnarize the studied signals. Most popular warping operators used in the literature are the functions $w(x) = \exp^x$ [1] and $w(x) = |x|^{1/k}\, sgn(x)$ that have shown themselves really usefull for respectively analyzing logarithm and monomial phase signals. They turn the signals into stationary pure sinusoids that are easier to analyze.

If the warping function is analytical and bijective, then the warped signal can be unwarped in order to express the signal in its original basis. However, they are most of the time not inversive as signals often met in real world applications are usually nonstationary and have time-frequency contents that cannot be described by analytical functions. Jarrot et al. [2] proposed a technique that extends the class of warping operators to discrete-time sequences and respects invertibility conditions.

2.1.2. *Discrete formulation*

The proposed concept starts from the straighforward definition for the sampled discrete-time warping operator:

$$\begin{aligned}(\mathcal{W}s[n])[m] &= s_w[m] \\ &= s\left(w_d\left(\tfrac{m}{M-1}\right)(N-1)T\right)\end{aligned} \tag{6}$$

where $w_d(u)$ is the normalized function of $w(t)$ defined by:

$$\begin{cases} w_d : [0,1] \rightarrow [0,1] \in \mathcal{C}^1 \\ w_d(0) = 0 \\ w_d(1) = 1 \\ dw_d(u)/du \geq 0 \\ w_d(u) = \frac{w((N-1)Tu)}{w((N-1)T)} \end{cases} \tag{7}$$

The main difficulty is now to estimate the missing values of the signal by using the general class of interpolators. The first step is to express the signal into a linear shift-invariant space V_ϕ [3] defined by the kernel ϕ with $V_\phi = \mathit{Span}(\{\phi(t-kT), k \in \mathbb{Z}\})$ [4], as follows:

$$s(t) = \sum_n c[n]\,\phi(t-nT) \tag{8}$$

We restrict ourselves to the case of exact interpolation where $s(nT) = s[n]$. Then, signal $s(t)$ can be expressed as:

$$s(t) = \sum_{n=0}^{N-1} s(n)\,\phi_{int}(t-nT) \tag{9}$$

where the kernel ϕ_{int} is:

$$\phi_{int}(t) = \sum_k p[k]\,\phi(t-kT) \tag{10}$$

Exact interpolation is possible if and only if the following condition is met for all signal:

$$\phi_{int}(nT) = \begin{cases} 1, & \text{if} \quad n = 0 \\ 0, & \text{if} \quad n \neq 0 \end{cases} \tag{11}$$

Note that any interpolation kernel can be used. Jarrot et al. [2] has found that the cardinal B-spline is the optimal choice. Therefore, the warped discrete sequence is defined as:

$$s_w[m] = \sum_{n=0}^{N-1} s[n]\,\phi_{int}(w_d(m) - nT) \tag{12}$$

More details on the algorithm is provided in [2]. This shows that any discrete time warping operator could be implemented thanks to a finite set of values of the warping function. This extends time axis transformations to non analytical warping functions which means that any kind of signals can be stationnarized by a sinusoid.

Let consider an arbitrary instantaneous frequency law $\psi(t)$ and a finite set of values $\{\psi(t_k)\}_{k=1,\ldots,L}$. The discrete warping function $w_d(t)$ that stationnarizes the signal is defined such as:

$$\begin{aligned}\psi(w_d(t_k)) &= w_d(\psi(t_k)) \\ &= t_k;\ k = 1,\ldots,L\end{aligned} \tag{13}$$

To solve this equation, the signal is divided into L segments in which a local estimation of the warping function is computed iteratively. Thus, the segmentation ensures $\psi(t)$ monotony on each interval. Each interval $[I_k, I_{k+1}]$ is centered around t_k and we note $w_k = w(t_k)$. The algorithm starts by initializing $w_k^{(0)}$ such as:

$$w_k^{(0)} = \frac{1}{2}(I_k + I_{k+1}) \tag{14}$$

Then, at each iteration p, the local solution is calculated as follows:

$$w_k^{(p)} = w_k^{(p-1)} - \mathit{sgn}\left(\psi\left(w_k^{(p-1)}\right) - t_k\right)\frac{I_{k+1} - I_k}{2^{p+1}} \tag{15}$$

where *sgn* is the sign function. The process stops when the differential value between two successive iterations becomes insignificant:

$$\left| w_k^{(p+1)} - w_k^{(p)} \right| < \epsilon \tag{16}$$

This process enables the stationnarization of any kind of instantaneous frequency laws even when non analytical.

2.2. Warping-based compressive sensing algorithm

Let consider the case of a composite discrete signal $s(n)$ composed of a desired component $s_d(n)$ that should be extracted and an undesired part $s_{nd}(n)$ that should be removed:

$$s(n) = s_d(n) + s_{nd}(n) \tag{17}$$

The desired signal corresponds here to a sinusoid. The other part can be composed of other coherent signals, as well as nonstationary or more complex signals. The goal is therefore to extract the sinusoid that can be well hidden into $s(n)$ and also more powerless than the undesired part.

Therefore, to extract the component of interest, it is necessary to explore the time-frequency content of the signal. We perform a short time Fourier transform (STFT) of $s(n)$ using a rectangular window of width M. At instant n, we have:

$$STFT(n,k) = \sum_{p=0}^{M-1} s(n+p)\, e^{-j2\pi kp/M} \tag{18}$$

Because the desired and undesired frequency component can overlap in time and frequency, it is then necessary to separate them. To do so, one possibility is to use the L-statistics approach [5, 6] to select time-frequency regions of interest.

Indeed the desired signal is sparse in the Fourier domain, so along a given frequency line, it is possible to have only undesired components or a mix of desired and undesired components. In both cases, the highest values of the spectrogram would correspond to interference contributions or unwanted components. Another scenario can be considered when the amplitude of the two parts's spectrum are of the same order of amplitude, it is known that opposite phases produce low values at the intersection points. All these informations lead to the assessment that we need to remove those values from consideration. To do so, data samples are first sorted out for each frequency. Then, the highest and the smallest values are discarded (i.e. replaced by a zero value). The procedure is described below.

The L-statistics is applied to $STFT(n,k)$. For each frequency k, the samples are sorted out:

$$L_k(p) = \mathit{sort}\left\{STFT(p,k),\quad p = 0,1,...,M-1\right\} \tag{19}$$

such that $|L_k(0)| \leq |L_k(1)| \leq ... \leq |L_k(M-1)|$. The removal operation is then done by discarding Q highest values and P smallest values of each L_k. The remaining STFT values consequently belong to the desired signal and a compressive sensing (CS) reconstruction algorithm can be used to reconstruct the sinusoid of interest.

At instant n, the STFT of signal $s(n)$ can also be expressed as follows:

$$\mathbf{STFT}_M(n) = \mathbf{D}_M \mathbf{s}(n) \tag{20}$$

where $\mathbf{STFT}_M(n)$ and $\mathbf{s}(n)$ are vectors in the form:

$$\mathbf{STFT}_M(n) = \left[STFT(n,0), ..., STFT(n,M-1)\right]^T \tag{21}$$

$$\mathbf{s}(n) = \left[s(n), s(n+1), ..., s(n+M)\right]^T \tag{22}$$

while $\mathbf{D}_M$ is the MxM DFT matrix with coefficients:

$$\mathbf{D}_M(m,k) = e^{-j2\pi mk/M}$$

As the case of nonoverlapping windows is considered, the complete formulation yields to:

$$\mathbf{STFT} = \mathbf{Ds} \tag{23}$$

such that vector $\mathbf{STFT}$ is composed of vectors: $\mathbf{STFT}_M(0)$, $\mathbf{STFT}_M(M)$, ..., $\mathbf{STFT}_M(N-M)$. The matrix $\mathbf{D}$ is obtained as a Kronecker product:

$$\mathbf{D} = \mathbf{I}_{N/M} \otimes \mathbf{D}_M$$

where $\mathbf{I}_{N/M}$ is the (N/M)x(N/M) identity matrix. The vector s can then be expressed as the inverse Fourier transform of the DFT vector $\mathbf{S}$:

$$\mathbf{s} = \mathbf{D}_N^{-1}\mathbf{S} \tag{24}$$

By combining (23) and (24), a general expression can be obtained:

$$\mathbf{STFT} = \mathbf{A}_{FULL}\mathbf{S} \tag{25}$$

with $\mathbf{A}_{FULL} = \mathbf{D}\mathbf{D}_N^{-1}$.

To summarize, matrix $\mathbf{A}_{FULL}$ maps the integrality of the TF representation. By using the L-statitics method [5], we create a CS mask $\mathbf{A}_{CS}$ that will remove the TF values that do not belong to the desired signal from the spectrogram. We then dispose of a spectrogram $\mathbf{STFT}_{CS}$ that has been cleansed of all undesired signals frequency components. The next step is to reconstruct the desired signal by minimizing the following problem:

$$\mathit{min}\,\|X\|_1 = \mathit{min} \sum_{k=0}^{N-1} |X(k)| \quad \textit{subject to} \tag{26}$$

$$STFT_{CS} = A_{CS}X$$

The reconstructed DFT $\mathbf{X}$ is used to obtain $s(n)$ which corresponds to $s_d(n)$ due to the elimination of all the non desired components.

To conclude, the compressive sensing part and the L-statistics approach manage to separate the sinusoid of interest from $s(n)$. Then to reconstruct the nonlinear modulation phase signals, it is necessary to perform an unwarping operation.

The Warping Based Compressive Sensing algorithm (WBCS) is the result of two contributions: the warping described in the previous section and a compressive sensing method introduced by Stankovic et al. [5]. It permits the extraction of nonlinear modulation phase signals even when they overlap in time and frequency, and also when parts of them are missing. First, the warping operation enables to express the desired signal into a pure sinusoid that is sparse in the Fourier domain, then, the CS algorithm combined with the L-statistics approach removes all the undesired components. Even if this approach seems similar to the combination of warping operators with band-pass filterings, it is bit a different as we will see in the next section. Warping operators defined by Jarrot et al. enable to estimate a well-suited warping function from a set of IFL coordinates and the compressive sensing permits to reconstruct a sinusoid even if parts of the signal are missing. This is this point that provides the main contribution of this work and enables to obtain better results compared to a more classical method.

3. NUMERICAL EXAMPLE

3.1. Presentation of the example

Let consider a signal $s(n)$ defined for $n \in \{1, 2, \ldots, N\}$, composed of two nonlinear modulation phase signals $s_1(n)$ and $s_2(n)$ that model mammals vocalizations:

$$s(s) = s_1(n) + s_2(n) \tag{27}$$

with:

$$s_1(n) = 2.5e^{j(795ln(n/N+1)+32n/N)} \tag{28}$$

$$s_2(n) = 1.5e^{j(1050ln(n/N+1)+100.8n/N)} \tag{29}$$

We simulate an echo of those signals (for the reflection with the sea's surface) by performing a convolution between $s(n)$ and the following transfer function $h(n)$ defined such as:

$$h(n) = \delta(n_1) + 0.7\delta(n_2) \tag{30}$$

with δ the dirac funtion, $n_1 = 10$ and $n_2 = 400$ the time instants when the direct path and the echo are respectively received by the microphone. The signal of observation $x(n)$ is as follows:

$$x(n) = s(n) * h(n) + b(n) \tag{31}$$

where $*$ stands for the convolution product and $b(n)$ for a white gaussian noise.

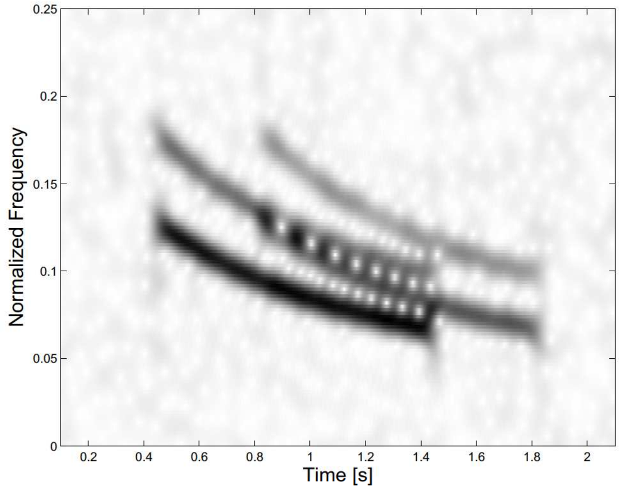

Fig. 1. Spectrogram of the numerical example obtained for a SNR equal to $7.95\ dB$.

In this paper, the signal-to-noise ratio (SNR) is defined as follows:

$$\mathbf{SNR}_{dB} = 10\log_{10}\frac{\sum_{n=1}^{N} s_d^2(n)}{\sum_{n=1}^{N} s_{nd}^2(n)} \tag{32}$$

with $s_{nd}(n)$ refering to unwanted signals such as the noise $b(n)$ and s_d the components of interest (direct path and echos).

Figure 1 presents the spectrogram of the studied signal using a 128 samples length Hamming window and 127 samples of overlapping. As we can see one component and an echo are so close in the TF plane that some artefacts are created and prevent tracking algorithms from providing a good estimation of the nonlinear modulations. In this example, we propose to extract $s_1(n)$ and $s_2(n)$ by using a warping approach.

3.2. Method

To begin with, we propose to work on a portion of $x(n)$ that only contains the two modulations of interest and the artefacts created by the echos. To extract the first component, we first estimate a set of its IFL's coordinates. This is performed by an algorithm of tracking phase developed by [7] which manages to track the time-frequency component based on their local continuity. To do so, the signal is divided into intervals in which an estimation of the local-best matched IFLs is performed. In this example, the algorithm provides a set of 32 coordinates for the IFL which is enough to estimate a warping function that stationnarizes the component.

Figure 2(a) shows the results of the tracking algorithm and the set of IFL coordinates estimated (b) used to calculate the associated warping function (c). After the warping operation, the signal's most powerfull component is expressed as a pure sinusoid mixed with perturbations in the warped domain. We

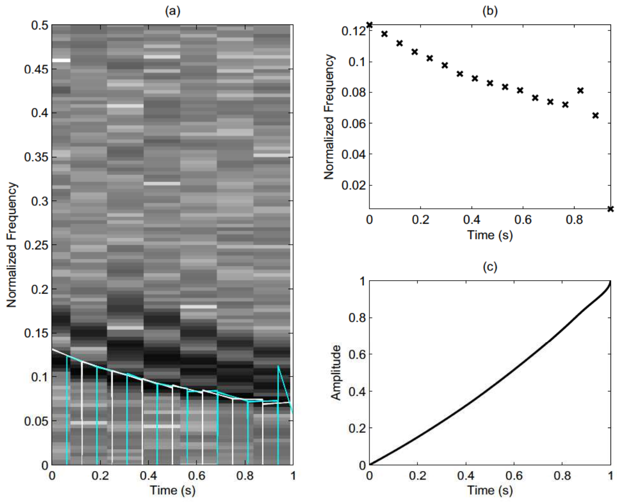

Fig. 2. (a) Spectrogram of the signal and result of the tracking method for the estimation of the IFL coordinates of the most powerfull signal (b). The associated warping function that stationnarizes the corresponding component is then estimated in (c).

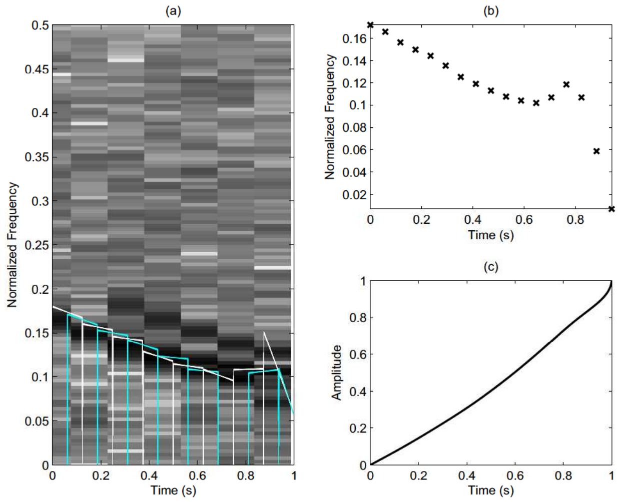

Fig. 4. (a) Spetrogram of the filtered signal and result of the tracking method for the estimation of the IFL coordinates of the second component (b). The associated warping function that stationnarizes the corresponding component is then estimated in (c).

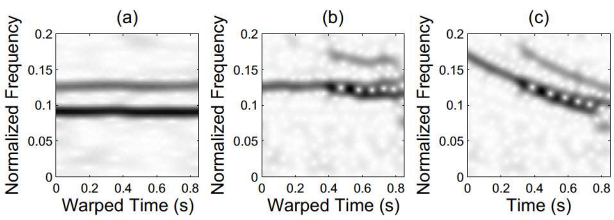

Fig. 3. (a) Spectrogram of the warped signal. (b) Spectrogram of the filtered warped signal: the most powerfull component has been removed. (c) Spectrogram of the filtered signal in the original time domain.

notice here that even if the IFL traking is subject to error, it enables to stationarize the component of interest.

The next step is then to use a band-pass filter to extract the component that has been stationnarized, perform an inverse time-axis transformation to express the selected component in the original time domain as $\widehat{s_1}(n)$ and finally calculate its IFL. The band-pass filter is a 300-th order filter that has a bandpass of $\frac{0.003}{N}$ centered around the central frequency given by the maximal peak of frequency of the warped signal's spectrum. Figure 3(a) shows the spectrogram of the warped signal where the most powerful component is represented by a sinusoid.

Afterwards, we perform a notch filter (with the same parameters used for the band-pass filter) to remove the first component from the warped signal (Figure 3(b)). The filtered signal is then expressed in the original time domain with an inverse time-axis transformation: we obtain the original signal reduced with the first component (Figure 3(c)). This step is necessary to estimate the second component as it becomes the most powerfull component and therefore can be trackable with the tracking algorithm. We then estimate a set of the second component IFL's coordinates in order to estimate a warping function that stationnarizes it and perform the associated warping operation (Figure 4). We then estimate a set of the second component IFL's coordinates in order to estimate a warping function that stationnarizes it and perform the associated warping operation (Figure 4). The second component is turned into a single sinusoid with missing values in the time-frequency plane (Figure 5).

The natural approach is then to perform a second band-pass filtering to extract the second component, followed by an inverse time-axis transformation. This is usually where the method becomes limited. Due to the missing values of the second component, the algorithm fails to extract properly the signal of interest $\widehat{s_2}(n)$. This is shown in Figure 7 where we can see that the estimated IFL of $\widehat{s_2}(n)$ is not perfect.

We then propose to observe the results obtained with the WBCS algorithm starting from the linearization of the second component. The STFT of the warped signal is then performed using a 32 samples rectangular window with no overlapping (Figure 6(a)). For each frequency, the STFT values are sorted out (Figure 6(b)) and according to the L-statistics process, 50% of the largest values are removed along with 8% of the smallest values. The CS spectrogram values that remain after the L-statistics removal are shown in Figure 6(c). A reconstruction algorithm is then applied to the CS STFT which permits to reconstruct the sinusoid of interest. Subsequently, an inverse time-axis transformation is performed to recontruct the nonlinear modulation phase signal $\widehat{s_2}(n)$ in the

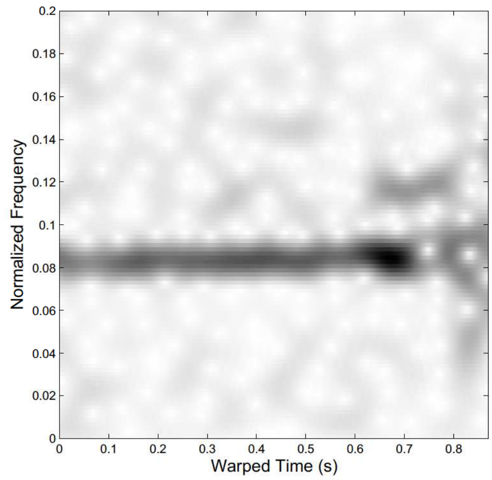

Fig. 5. Spectrogram of the second component after the warping process.

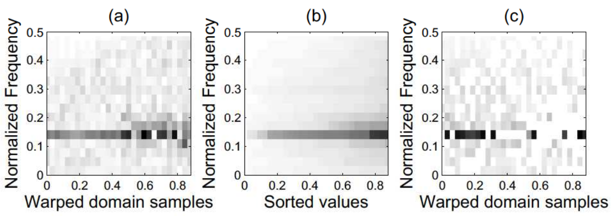

Fig. 6. (a) STFT of the warped signal, the sorted values for each frequency of the STFT (b) and the CS STFT remaining after the L-statistics process (c).

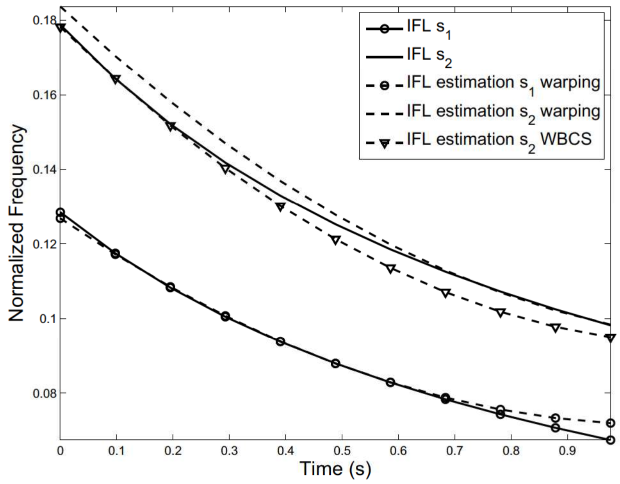

Fig. 7. IFL estimates of $s_1\,(n)$ and $s_2\,(n)$ calculated from the original signals, and the extracted signals obtained with the warping and WBCS algorithms.

original time domain.

Figure 7 shows the IFL of $\widehat{s_2}\,(n)$ obtained with the WBCS algorithm. As we can see, the result is better than the one obtained with the traditionnal band-pass filtering and enables to reconstruct the signal of interest even if a part is missing.

3.3. Results presentation

As we can see, results highly depend on the IFL estimation that enables to linearize the nonlinear component and they can be affected by the noise level. This is why, for a given level of noise that vary around $7dB$, we propose to run this experiment $N_{iter} = 300$ times and calculate the associated mean square error (MSE) to verify that the WBCS method provides a better accuracy than the combination of warping operators and band-pass filterings. The MSE is calculated as follows:

$$\textbf{MSE} = \frac{1}{N} \sum_{k=1}^{N} (y_k - \widehat{y_k})^2 \tag{33}$$

where y_k stands for the known IFL and $\widehat{y_k}$ for its estimate.

Results show that for only 3 cases both methods fail to provide a good estimation of the second IFL. The mean values of the MSE are respectively of 29.55 Hz and 24.09 Hz for the first method and the WBSC algorithm. In both cases, the accuracy of the result depends on the IFL estimation made for the warping process. The advantage comes from the possibility to reconstruct signals when parts of the signal are missing.

4. CONCLUSIONS AND PERSPECTIVES

This article proposed a new approach for nonlinear frequency extraction that enables to overcome the limitation when parts of the signals are missing. It takes advantages of a warping transformation that turns the desired signal into a sinusoid and the L-statistics that permits to select time-frequency regions of interest. Then, a CS reconstruction algorithm is used to regenerate the signal in the warped domain. Finally, a second time axis transformation is performed to express the reconstructed warped signal in the original time domain.

This work is supported by the Montenegrin Ministry of Science, project grant "New ICT Compressive sensing based trends applied to: multimedia, biomedicine and communications" (ACRONYM: CS-ICT).

5. REFERENCES

[1] A. Papandreou-Suppappola, "Generalized time-shift covariant quadratic time-frequency representations with arbitrary group delays," in *Conference Record of The Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*. 1995, vol. 1, pp. 553–557, IEEE Comput. Soc. Press.

[2] A. Jarrot, C. Ioana, and A. Quinquis, "An Extension of the Class Of Unitary TimeWarping Projectors To DiscreteTime Sequences," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2006, vol. 3, pp. 412–415.

[3] E. Meijering, "A chronology of interpolation: from ancient astronomy to modern signal and image processing," *Proceedings of the IEEE*, vol. 90, no. 3, pp. 319–342, Mar. 2002.

[4] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited.," *IEEE transactions on medical imaging*, vol. 19, no. 7, pp. 739–58, July 2000.

[5] L. Stanković, I. Orovic, S. Stanković, and M. Amin, "Compressive Sensing Based Separation of Nonstationary and Stationary Signals Overlapping in Time-Frequency," *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4562–4572, 2013.

[6] L. Stanković, S. Stanković, and M. Amin, "Missing samples analysis in signals for applications to L-estimation and compressive sensing," *Signal Processing*, vol. 94, pp. 401–408, Jan. 2014.

[7] C. Ioana, C. Gervaise, Y. Stéphan, and J. I. Mars, "Analysis of underwater mammal vocalisations using timefrequency-phase tracker," *Applied Acoustics*, vol. 71, no. 11, pp. 1070–1080, Nov. 2010.