TIME-FREQUENCY TRACKING USING MULTI-WINDOW LOCAL PHASE ANALYSIS

Florian Dadouchi[1], James W. Pitton[2], Cornel Ioana[1], Cédric Gervaise[1]

[1]: GIPSA-Lab, Grenoble Institute of Technology,
38402, Saint Martin d'Heres – FRANCE
E-mail: [Florian.Dadouchi, Cornel.Ioana, Cedric.Gervaise]@gipsa-lab.grenoble-inp.fr
[2]: Applied Physics Laboratory, University of Washington
1013 NE 40th Street Box Seattle, WA 98105-6698 – USA
E-mail: pitton@apl.washington.edu

ABSTRACT

The analysis of signals consisting of multiple components with non-linear frequency modulation is required in a large number of applications, including the study of marine mammals vocalizations. This analysis has multiple motivations, such as investigating the impact of anthropogenic noise on marine mammal behavior, and species identification to avoid collision between ships and marine mammals. Such applications are normally conducted in a Passive Acoustics Monitoring (PAM) context, where there is low SNR together with very little a priori knowledge on the signals being analyzed. Recently, time-frequency tracking based on local analysis of the instantaneous phase has been successfully applied to underwater signals. In this paper, we present a robust version of this method, based on the use of the multiple analysis windows. The results provided for simulated and real signals demonstrate improved tracking of the instantaneous frequency in noise.

Index Terms - *Time-Frequency Tracking, Instantaneous Frequency, Multi-window.*

1. INTRODUCTION

The significant increase in coastal human activities in recent years has resulted in a dramatic rise of underwater noise pollution and collisions between ships and marine mammals [1]. There is then a great issue at stake in being able to detect, classify, and localize marine mammal species through their acoustic behavior, *i.e.* their vocalizations. Dealing with large amounts of data and very little a priori knowledge of the observed signals, there is a tremendous need for the automated processing of marine mammal vocalizations. Since these signals are highly nonstationary, processing must center on estimation of the instantaneous frequencies of the signal components. Most of the current methods tackling this issue are spectrogram-based methods [2] suffering the classic limitations of time-frequency representations, especially when dealing with highly non-stationary signals characterized by non-linear frequency modulation. Recently, a time-frequency tracking methodology has been developed [3], exploiting the coherence of the fundamental parameters of a signal, namely the instantaneous amplitude, frequency and phase. Specifically, the time-frequency components are analyzed by local phase analysis, using polynomial phase modeling. Continuity between signal components is preserved locally through maximization of the local correlation between the signal's components defined by local polynomials.

In order to improve this technique with respect to noise robustness, we proposed a multi-window approach. This concept, introduced in [4] and revisited recently in the case of other time-frequency distributions [5], is adapted here to local polynomial modeling of the signal's components. Local polynomial modeling is performed in two windows of different durations and this operation simultaneously ensures good noise robustness (thanks to the largest window) and accurate phase modeling (thanks to the shortest window).

The paper is structured as follows. In section 2 we review tracking by local phase analysis. In section 3, the multi-window approach is defined in the context of local phase analysis. Results for both simulated and real underwater data are discussed in the section 4. We conclude in section 5.

2. TRACKING BY LOCAL PHASE ANALYSIS

Local phase analysis consists of ***second order local phase modelling,*** applied in local time windows that are half-overlapped. For each window, the 2nd order Warped HAF (High Order Ambiguity Function) is applied, providing 2nd order phase modelling of one component [3]. The WHAF of order 2 is defined as the integral of the HAF,

$$WHAF_k(\omega) = \int HAF_k[\omega, \tau_w]\, d\tau \qquad (1)$$

computed for the warped set of lags

$$\tau_w = \tau^{\frac{1}{k-1}} \qquad (2)$$

which comes from the dependence between the polynomial coefficient of order k, a_k, and the frequency corresponding the maxima of HAF:

$$\omega_k = k!\tau^{h-1}a_k \tag{3}$$

Using the warped set of lags, the peaks of the HAFs will be located on a line defined by $\omega_k = k!a_k$. Integrating the HAFs along this line will provide a more robust representation than an individual HAF.

In our example, naturally, WHAF estimation is applied to the most energetic component of the signal. The instantaneous frequency law (IFL) estimates obtained for two (overlapped) neighbour windows, *i* and *i+1*, are plotted in the figure 1. We note that the IFLs are close to the modelled component. Indeed, in real situations, this modelling is affected by noise and/or by nearby components. For this reason, we use WHAF to provide, for each window, several estimates of the same component (five in our simulations, which provide a satisfactory trade-off between tracking accuracy and computing time).

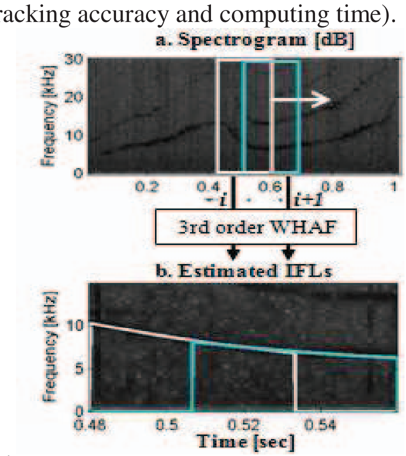

Figure 1. 3rd order local phase modelling using WHAF is applied in overlapped windows

For each window, WHAF provides N_c estimates of the most energetic time-frequency component. Let us denote the set of phase functions obtained from WHAF-based phase modelling, applied in i^{th} window, as:

$$D^{(i)} = \left\{\psi_k^{(i)}\right\}_{k=1,..,Nc} \tag{4}$$

where $\psi_k^{(i)}$ is the k^{th} phase function defined as:

$$\psi_k^{(i)}(t) = \sum_{i=1}^{2} a_{ki}t^i ; t \in \left[iT, (i+3/2)T\right] \tag{5}$$

where T is the window size.

The phase functions (or the IFLs) provided by the local polynomial phase analysis (4) are used to build **local filter functions** that extract the portion of the signal corresponding to the time-frequency regions defined in the neighbourhood of the local functions. This time-frequency filter is implemented as follows. The inverse function of the linear FM (frequency modulation) (5) is used to demodulate (*warp*) the portion of the signal corresponding to the time-frequency region defined by this linear FM. Once the corresponding component has been demodulated, a band pass filter is applied. The extracted component is then transformed back to the original time-frequency region by applying the inverse transformation (*unwarping*) using the considered quadratic FM. This filtering operation is done for all linear FMs provided by the first step.

Let consider the analyzed signal x, and two analysis windows *i* and *i+1*. Using the phase functions estimated in these windows, we construct two sets of time-frequency filters $\left\{\mathbf{W}_k^{(i)}\right\}_{k1,..N_c}$ and $\left\{\mathbf{W}_k^{(i+1)}\right\}_{k1,..N_c}$. Using these filters, we extract, from the analyzed signal x, the signal's time-frequency components denoted by:

$$\left\{s_k^{(i)}\right\} = \left\{\left(\mathbf{W}_k^{(i)}x\right)(t)\right\}_{k1,..N_c} ; \left\{s_k^{(i+1)}\right\} = \left\{\left(\mathbf{W}_k^{(i+1)}x\right)(t)\right\}_{k1,..N_c} \tag{6}$$

Using these extracted signal components, two phase functions $\psi_m^{(i)}$ and $\psi_n^{(i+1)}$ are assigned to the same time-frequency trajectory if the correlation of the corresponding components $\left\{s_m^{(i)}\right\}$ et $\left\{s_n^{(i+1)}\right\}; m,n = 1,...,Nc$ is maximal for all pairs (m,n):

$$\begin{aligned}(\hat{m},\hat{n}) = \max_{\substack{m\in[1,Nc]\\ n\in[1,Nc]}} \left|\left\langle s_m^{(i)}(t), s_n^{(i+1)}(t)\right\rangle\right| \Rightarrow \\ \Rightarrow \psi_{\hat{m}}^{(i)} \circ \psi_{\hat{n}}^{(i+1)} \subset \text{Trajectory TF } \phi_j\left[(i-1)T:(i+1)T\right] \text{ of signal } x\end{aligned} \tag{7}$$

where « ° » symbolizes the fusion of two phase functions belonging to the same trajectory defined on the time interval $(i\text{-}1)T\text{:}(i\text{+}1)T$. Actually, correlating two signals with close time-frequency content and overlapped time support allows us to compare their phases without having to estimate them. After analyzing all segments of the signal, we define the time-frequency trajectory as the fusion of phase functions following (7).

3. MULTI-WINDOW APPROACH

The previously described approach estimates the local polynomial coefficients in a given window. However, it is impossible to know *a priori* if the polynomial model of order 2 is optimal with respect to the window size. We therefore adopt a multi-window analysis approach taking advantage of the properties of various window lengths. Naturally, a short window will provide more accurate (limited bias) local phase modeling , but the reduced number of samples results in a lack of robustness to additive noise (large variance). In contrast, a longer window will improve the robustness (small variance), but the polynomial model of order 2 might not be sufficient to fit rapid variations of the IFLs (large bias). Our contribution consists in the definition of an automatic procedure for the choice of the size of the analyzing window.

Local second order phase modeling is equivalent to searching for the polynomial coefficients a_1 and a_2 (respectively, the center frequency and the chirp rate) in a given space, defined by the window size, T_1 and the bandwidth $[0,F_S/2]$ – F_s is the sampling frequency. Figure 2 illustrates the search space of (a_1,a_2).

The multi-window approach starts with estimation in a long-duration window. This estimate, accordingly to Fig. 2, is expected to be a rough approximation of the true component to estimate. Based on the robustness of WHAF, we expect the estimation error to be moderate.

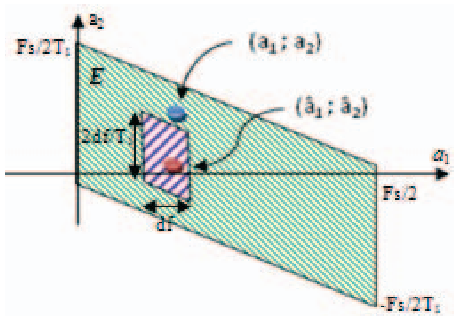

Figure 2. Search space for the vector (a_1,a_2) , the green shadow area is the search space for long-duration window

With this in mind, we reduce the search space for the component to the neighborhood of this first estimate. The search space is then defined by:

$$(\hat{a}_1,\hat{a}_2)\in E, \forall t\in\left[-\frac{T_1}{2},\frac{T_1}{2}\right] \quad (8)$$
$$f(t)=\hat{a}_1+2\hat{a}_2 t\in\left[f_{down}(t), f_{up}(t)\right]$$

where $(\hat{a}_1,\hat{a}_2)$ are the estimated values given by the WHAF in the window of size T_i and $f_{down}(t)$ and $f_{up}(t)$ respectively are defined by

$$f_{\substack{up\\down}}(t)=\hat{a}_1\pm\frac{df}{2}+2\hat{a}_2 t \quad (9)$$

where df will be chosen as the spread of the WHAF of first order. Other choices could be based on the minimization of the root square mean error or the maximization of the concentration in the time-frequency domain [5]. The goal of this first estimate is that, by using a long window, we obtain a *robust estimate* of the center frequency in the interval $\left[\hat{a}_1-\frac{df}{2},\hat{a}_1+\frac{df}{2}\right]$.

Once this interval is determined, the next step consists of WHAF application in a shorter window, of size T_2. The purpose of this new estimation stage is to improve the accuracy of chirp rate estimation. As indicated in the figure 3, analysis in a shorter window is equivalent to the dilation of the search space along the a_2 axis. The search space is then defined by:

$$(\hat{a}_1,\hat{a}_2)\in E, \forall t\in\left[-\frac{T_2}{2},\frac{T_2}{2}\right] \quad (10)$$
$$f(t)=\hat{a}_1+2\hat{a}_2 t\in\left[f_{down}(t), f_{up}(t)\right]$$

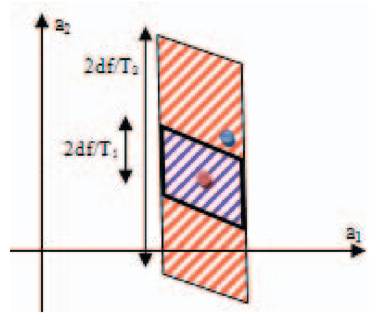

Figure 3. Dilated Search space for the vector (a_1,a_2)

This definition of the search space is very useful to eliminate incorrect estimates from the WHAF. As indicated in the previous section, the local analysis consists of short-time polynomial modeling of order 2, provided by WHAF. The result of this step is described by (4) and (5) and consists of several local chirps. Because of noise and interference, these local chirps could constitute an inaccurate estimation of the true components. For this reason, the multi-window approach defined in this section eliminates the WHAF estimated values that are outside the space defined by (8) and (10). In this way, we avoid these incorrect components, as we will see in the next section.

4. RESULTS

For the first example, we consider a synthetic sinusoidal frequency modulation (SFM) embedded in white Gaussian noise with SNR of -1.5 dB.

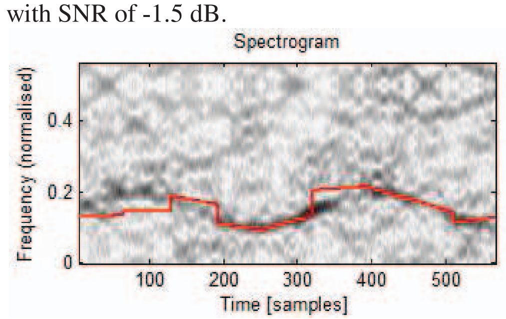

Figure 4. Tracking of a synthetic sinusoidal frequency modulation using a single window of 128 points

Figure 4 shows the tracking result using local phase analysis and the fusion of local information defined by (7). As the figure shows, the tracking result has some errors due to the noise components. This is visible at the beginning where the errors of the chirp rate estimation result from, to first order, an inaccurate estimate of the center frequency.

The solution based on multi-window approach shows its potential in the next two figures. Figure 5 corresponds to tracking with a window of size 192 points. Compared with figure 4, we observe that the center frequencies (a_1 coefficients) are better estimated in all windows. The estimated coefficients are used to generate the new search space (10) with a window size of 128 points. Figure 6 illustrates the tracking result obtained by the local analysis in the search space defined by (10). We remark that the tracking is accurate in both chirp rate and center frequency. This example illustrates the improvement provided by the multi-window approach with respect to the single-window approach described in section 2.

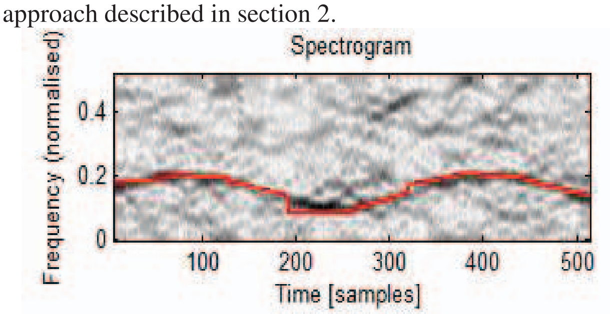

Figure 5. Tracking based on local phase analysis in a window of 192 samples

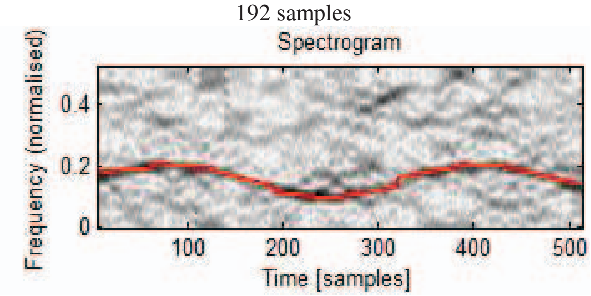

Figure 6. Tracking based on local phase analysis with 2 windows algorithm (window$_1$: 192 samples, window$_2$: 128 samples)

The next example concerns the tracking of a real marine mammal vocalization. Figure 7a indicates the tracking result for the single window approach using a 1024 point window. As indicated by the circles, the tracking fails to preserve the continuity of the time-frequency component.

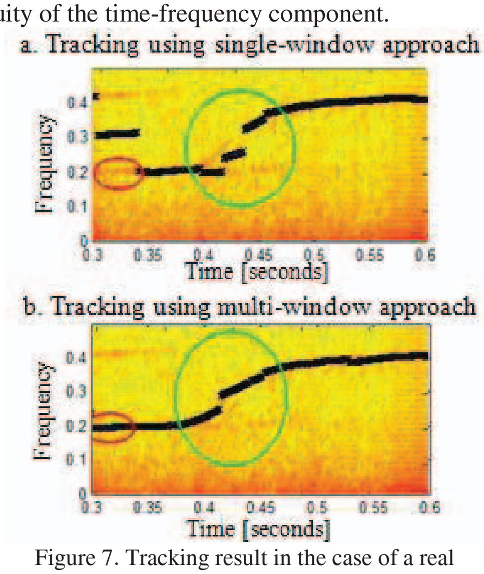

Figure 7. Tracking result in the case of a real marine mammal vocalization (a : single window time-frequency-phase tracking, window length = 1024 samples, b : two windows time-frequency-phase tracking window 1 length = 1024, window 2 length = 128

The multi-window approach solves this problem, as illustrated in figure 7b. Local polynomial phase modeling in the reduced search space gives a more accurate estimation of local components. The global tracking result is then better than the result from a single-window approach.

5. CONCLUSIONS

In this paper we defined a multi-window tracking approach in the context of time-frequency tracking of complex signals characterized by non-linear frequency modulated components. The fundamental concept underlying the method is time-frequency-phase coherence, which is locally exploited by polynomial phase modelling of order 2. In order to improve the estimation accuracy and noise robustness, a multi-window approach is used. First, the local phase analysis is done in a long-duration window, and then a shorter window is used in order to improve the estimation of the chirp rate.

The results, provided for synthetic and real data, proved the efficacy of the proposed methodology. Although the improved performance is illustrated in an underwater application, the approach presented here is more general since it exploits fundamental aspects of many types of signals – namely, the continuity of the instantaneous phase.

In future work, we will concentrate both on theoretical and applicative aspects. Concerning theoretical work, the extension to higher order polynomial modelling will be investigated. Subsequent versions of this methodology will be studied in applications such as signal classification and localisation.

Acknowledgment. This work has been supported, in part, by the French DGA project “GREENAR”.

5. REFERENCES

[1] T. J. O’Shea, D. K. Odell, “ Large-scale marine ecosystem change and the conservation of marine mammals”, Journal of Mammalogy, 89(3):529–533, 2008.

[2] T. A. Lampert, S E.M. O’Keefe “A survey of spectrogram track detection algorithms”, Journal of Applied Acoustics 71 (2010) 87–100.

[3] C. Ioana, J. Mars, A. Serbanescu, S. Stankovic, “Time-frequency-phase tracking approach: application to underwater signals in passive context”, Special Session on “Time-Frequency analysis in underwater applications”, *IEEE ICASSP*, Dallas, March 2010.

[4] G. Frazer, B. Boashash, “Multiple window spectrogram and time-frequency distributions,” in *Proc. IEEE ICASSP*, vol. 4, 1994, pp. 193-296.

[5] I. Orovic, S. Stankovic, T. Thayaparan, LJ. Stankovic, “Multiwindow S-method for Instantaneous Frequency Estimation and its Application in Radar Signal Analysis,” *IET Signal Processing*, Vol. 4, No. 4, pp: 363-370, Jan. 2010.