Parameters estimation for non-linear resynthesis methods with the help of a time-frequency analysis of natural sounds.


Abstract: The goal of our study is to build a real-sound analysis method which allows to estimate parameters for non-linear synthesis techniques such as Frequency Modulation, Amplitude and Phase Modulation. Using the notion of "ridge" associated with certain kinds of time-frequency analysis (Gabor or wavelet transforms), we shall describe how such trajectories can lead to modulation laws. Estimation of non-linear synthesis parameters such as A.M. and P.M. has been performed on instrumental sounds. In this case, several trajectories are generally extracted, leading to a combination of additive and non-linear synthesis methods.

1 Introduction

Our purpose is the establishment of a correspondence between algorithmic synthesis techniques (Amplitude Modulation (A.M), Frequency Modulation (F.M), Phase Modulation(P.M)), and particular trajectories in the time frequency plane, which represent modulation laws of real signals. These trajectories can be interpreted as the "score" associated to the sound (fig.1,2,3).

![Fig 1: Sine wave](image1)

![Fig 2: Linear chirp](image2)

![Fig 3: Frequency modulation](image3)

In the next sections, we shall first explain how to determine the frequency modulation trajectories in the time-frequency representation. A very simple algorithm will be described and results obtained with synthetic signals will be presented.

The estimation of non-linear synthesis parameters such as Amplitude Modulation and Phase Modulation has been performed on natural sounds. The estimation of Frequency Modulation parameters is still in progress: we shall only describe the importance of the choice of some analysis parameters such as the window-size in the Gabor decomposition and the Q-factor in the wavelet transform.

2 How to determine the modulation trajectories

2.1 Global trajectories

It is always possible to associate to a signal $s(t)=A(t)\cos(\phi(t))$ the analytic signal $z(t)=b(t)e^{j\psi(t)}$ (the imaginary part of $z(t)$ is given by the Hilbert transform of $s(t)$) [1]. Generally, $A(t)=b(t)$, $\phi(t)=\psi(t)$ and the representation obtained is difficult to modulate since the informations about $A(t)$ and $\phi(t)$ are dispatched at both $b(t)$ and $\psi(t)$.

These problems are less severe if this kind of decomposition is done in limited frequency ranges.

2.2 Local trajectories

2.2.1 Time-and-scale, time-frequency decompositions.

The aim of such a decomposition is to associate to a signal $s(t)$ a set of coefficients labelled with two parameters: one related to time and one related to frequency. This operation leads to a transform which pictures both the time and the frequency variations of the signal. Some examples of decompositions are:

ICMC GLASGOW 1990 PROCEEDINGS

88
Gabor: \( \psi(t) \rightarrow W(\lambda_0) \) time-frequency representation
Wavelet: \( \psi(t) \rightarrow S(\lambda_0) \) time-scale representation. \( \lambda_0 \) is the scale or dilation parameter.
For more information see e.g. [2].
These operations on a signal introduce redundancy on the data, and it is natural to think that all the coefficients have not the same "weight".
The next section discusses how to extract a "ridge" in the time-frequency plane of the transform, corresponding to the more important values.

2.2.2 Ridge extraction.
We have shown that the main contributions to the transform are given by stationary points [3]. A practical criteria for their identification is:

\[ \frac{d\phi(\lambda_0)}{d\lambda_0} = 0 \]

Wavelet: \( \frac{d\phi(\lambda_0)}{d\lambda_0} = 0 \)

where \( \phi \) represents the phase of the wavelet, \( \lambda_0 \) the frequency associated to the Gabor representation, \( \lambda_0 \) the dilation or scale parameter, \( \frac{d\phi}{d\lambda_0} \) the frequency of the analyzing wavelet at the scale \( \lambda_0 \).
It can be shown [4] that these points define a modulation trajectory in the plane of the transform. Moreover, the phase and the modulus of the transform along one of these trajectories give a good estimation of the phase and amplitude of the signal in a limited frequency band.

2.2.3 Examples
Figure 4 represents from top to bottom the waveform, the ridge and the modulus of the wavelet transform of an hyperbolic chirp.
Figure 5 represents the waveform, the ridge and the modulus of the wavelet transform of a F.M signal with a time increasing index.

Fig 4: hyperbolic chirp
Fig 5: F.M. with linear index

3 Interpretation of the modulation trajectories
In a general way, catching modulation laws depends mainly on the choice of the Q-factor of the wavelet or on the size of the Gabor's frequency window with respect to the frequency band of the analysed signal.
- If the window is too narrow, one catches each harmonic separately just as a sonagram does.
- If the window is too large, we have the same problems as for the global analysis described above.

In figure 4, the extracted frequency modulation law exactly corresponds to the theoretical parameters.
Figure 5 shows what happens with a complex synthetic sound. The signal analyzed with the help of the wavelet transform is a F.M. signal, with linear time-dependent index.
We can see that several ridges are extracted. One can roughly explain that:
- At low frequencies the wavelet is very narrow in frequency so we catch individual components or beats between two close components.
- At high frequencies we extract frequency modulation laws. In fact these modulations do not correspond exactly to the ones needed for resynthesis. This is due to low frequencies components which are caught independently. Nevertheless, the shape of the modulation index is respected and we get exact values for carrier and modulation frequencies.

ICMC: GLASGOW (199) PROCEEDINGS
4 Analysis-resynthesis of musical sounds

The study on academic F.M. and P.M. signals has helped us to interpret the ridge and to understand the relation between the Q-factor of the wavelet and the modulation laws extracted.

We are currently testing what happens with musical sounds and studying the resynthesis from the extracted trajectories, a very interesting step.

4.1 Synthesis from the ridges

We have done a resynthesis combining additive and non-linear methods (Amplitude modulation and Phase Modulation) (Fig.6), in the case of complex sounds we have several trajectories. For each trajectory, we extract the phase modulation law and the modulus as described above. The summation of each contribution (for each trajectory) gives the signal. This has been done on a saxophone sound (Fig.7) and a trumpet sound with five trajectories; the resynthesis are quite satisfying.

![Fig 6: Resynthesis algorithm](image)

![Fig 7: Wavelet transform: 250 ms of Saxophone sound](image)

From top to bottom: waveform, phase and modulus as a function of time

4.2 Conclusion and next objective

Our ultimate purpose is to estimate frequency modulation parameters and not simply a non-periodic or irregular function (as in the case of F.M.). In this case it is necessary to continuously adapt the size of the Cohen's window (or the Q-factor for wavelets).

We have just described an approach of the problem and we are now working on the estimation of the right synthesis parameters for FM synthesis.

References


ICMC GLASGOW 100 PROCEEDINGS