An Entropy Based Method for Local Time-adaptation of the Spectrogram

Marco Liuni[*][**], Axel Röbel[**], Marco Romito[*], and Xavier Rodet[**]

[*]Università di Firenze, Dip. di Matematica "U. Dini"
Viale Morgagni, 67/a - 50134 Florence - ITALY
[**] IRCAM - CNRS STMS, Analysis/Synthesis Team
1, place Igor-Stravinsky - 75004 Paris - FRANCE

This work is supported by grants from Region Ile-de-France

{marco.liuni,axel.roebel,xavier.rodet}@ircam.fr
marco.romito@math.unifi.it
http://www.ircam.fr/anasyn.html

Abstract. We describe a method for local time-adaptation of the spectrogram of audio signals: it is based on the decomposition of a signal within a Gabor multi-frame of raised cosine windows through the STFT operator. The sparsity of the analyses in every single frame of the multi-frame is evaluated through the Rényi entropies measures.
We give an analytical expression for the entropy of the spectrogram of a sinusoidal signal using Hanning windows of different sizes, and describe the results obtained with other basic signals; we then provide an example of the performance of our algorithm with an instrumental sound.

Key words: adaptive spectrogram, sound representation, sound analysis, sound synthesis, Rényi entropies, sparsity measures, frame theory.

1 Introduction

Far from being restricted to entertainment, sound processing techniques are required in many different domains: they find applications in medical sciences, security instruments, communications. The most challenging class of signals to consider is indeed music: the completely new perspective opened by contemporary music with the deep reconsideration of concepts as noise and timbre makes every sound a potentially musical sound.

The standard techniques of digital analysis are based on the decomposition of the signal in a system of elementary functions, and the choice of a specific system necessarily has an influence on the result. This motivates the search for adaptive methods of sound analysis and synthesis, and for algorithms whose parameters are designed to change according to the analyzed signal features. Our research is focused on the development of mathematical models and tools based on the local automatic adaptation of the system of functions used for the decomposition of the signal: we are interested in a complete framework for analysis, spectral transformation and re-synthesis; thus we need to define an efficient strategy to reconstruct the signal through the adapted decomposition, which must give a perfect recovery of the input if no transformation is applied.

Here we propose a method for local automatic time-adaptation of the Short Time Fourier Transform window function, through a minimization of the Rényi entropy of the spectrogram; we then define a re-synthesis technique with an extension of the method proposed in [9]. Some examples of this approach can be found in the literature: the idea of gathering a sparsity measure from Rényi entropies is detailed in [1], and in [11] a local time-frequency adaptive framework is presented exploiting this concept, even if no methods for perfect reconstruction are provided. The choice of Frame Theory ([2],[10]) as a model for the analysis/synthesis process is largely described in [14]: the fundamental step to Multiple Gabor Frames is comprehensively treated in [6] and an approach where sparsity is obtained through a regression model is introduced in [16]; a recent development in this sense is contained in [12] where a class of methods for analysis adaptation are obtained separately in the time and frequency dimension together with perfect reconstruction formulas: indeed no strategies for automatization are employed, and adaptation has to be managed by the user. The model conceived in [15] belongs to this same class but presents several novelties in the construction of the Gabor multi-frame and in the method for automatic local time-adaptation. In [13] another time-frequency adaptive spectrogram is defined considering a sparsity measure called *energy smearing*, without taking into account the re-synthesis task. The concept of *quilted frame*, recently introduced in [7], is the first promising effort to establish a unified mathematical model for all the various frameworks cited above.

In the second section we recall the concept of spectrogram and the link between the analysis resolution and the size of the window used. In the third we apply the method introduced in [1] to obtain a global information about the sparsity of a time-frequency density using the Rényi entropies, providing new analytical results for the class of spectrograms with raised cosine windows. The fourth section describes how to model the problem in the Frame Theory domain, as the decomposition of a signal within a Gabor multi-frame, and we use this approach to derive a local entropy measure which rules the choice of a local optimal resolution. In the fifth section we finally provide a description of the algorithm and an example of adapted spectrogram for a monophonic instrumental sound.

2 Spectrogram as Energy Density

Several works have investigated the relations between the physical and the probabilistic concepts of density (see [3],[4]); in the field of time-frequency analysis, the interest is focused on functions $\Phi \in L^2(\mathbb{R}^2)$ which jointly represent the energy of a function $f \in L^2(\mathbb{R})$ and the energy of its Fourier transform. This approach leads to consider a large class of representations of a signal as time-frequency energy densities, and to the use of probabilistic tools to analyze their features.

A *window function* is a function $g \in L^2(\mathbb{R})$ such that $\|g\| = 1$ and $g(t) = g(-t)$. The related transform operator is called *Short Time Fourier Transform (STFT)* and it is defined as

$$\mathrm{S}f(u,\xi) = \langle f, g_{u,\xi} \rangle = \int f(t) g(t-u) e^{-2\pi i \xi t} \mathrm{d}t, \quad \forall f \in L^2(\mathbb{R}). \tag{1}$$

The *spectrogram* is the squared modulus of the STFT

$$\mathrm{PS}_f(u,\xi) = |\mathrm{S}f(u,\xi)|^2 = \left| \int f(t) g(t-u) e^{-2\pi i \xi t} \mathrm{d}t \right|^2 , \tag{2}$$

and it measures the signal energy in a neighborhood of (u, ξ) in the time-frequency plane. The spectrogram can be considered as a joint time-frequency density of the energy of the signal f, whose time and frequency marginals are respectively $|f(t)|^2$ and $|\hat{f}(\omega)|^2$. Actually, the marginals are not satisfied as they depend on the window function, nevertheless this does not affect the advantages of considering the spectrogram as a density, with cares to the necessary distinctions.

2.1 Spectrogram Resolution

The resolution of analyses based on time-frequency transforms, such as the STFT and the Wavelet Transform ([8],[14]), is linked to the time and frequency concentration of the basic functions involved in the decomposition, represented by their *Heisenberg boxes* (for a definition see [14] chapter IV): these are rectangles drawn in the time-frequency plane whose dimensions are linked respectively to the time spread of a function and to the frequency spread of its Fourier Transform. In the STFT, the boxes associated to the transpositions of the window function have fixed dimensions in every area of the time-frequency plane: the resolution is the same for all the components of the signal.

The analysis resolution can be globally modified with the choice of a different window or by a scaling operation

$$g^l(t) = \frac{1}{\sqrt{l}}\, g\left(\frac{t}{l}\right) , \tag{3}$$

which has the effect of changing the ratio between the edges of the Heisenberg box associated to g preserving its area: this means that we are changing the resolution by privileging the concentration in one dimension to the detriment of the other.

There are limits to the achievable analysis accuracy imposed by the *Heisenberg uncertainty principle* ([4] chapter III, [14] chapter II): its interpretation is that more accuracy in the time domain resolution causes a loss of precision in the frequency domain, and vice versa. A poor time resolution can lead to uncertainty in the time location of an event, i.e. the event is not detected exactly when it happens but at several times around its true location; a poor frequency resolution can lead to representations which are not consistent with the perceptive features of the signal: this happens for example when two different spectral components are so close in frequency that they are merged together in the spectrogram, and so not resolved.

A method for a local adaptation of the time-frequency resolution gives a more flexible and coherent representation of the signal, in particular providing a higher precision for the existing techniques of processing: we envisage that an adaptive analysis and re-synthesis framework would give better performances within all the algorithms based on spectral processing and sinusoidal models. Nevertheless, the definition of the optimal tradeoff between time and frequency resolution is not unique; several criteria exist to classify the *sparsity* of the analysis of a signal, i.e. the concentration of the energy density given by a certain time-frequency representation of the signal. In our method, we use the sparsity measures introduced in [1] which are based on Rényi entropies, as we describe in the next section.

3 Rényi Entropies

Thanks to the above interpretation of the spectrogram as an energy density, some techniques belonging to the domain of Probability and Information Theory can be applied to our problem. In particular, the concept of *entropy* can be extended to give a sparsity measure of a time-frequency density. A promising approach ([1]) takes into account Rényi entropies, a generalization of the Shannon entropy: they are defined for an *order* $\alpha > 0,\ \alpha \neq 1$ and for a time-frequency representation $\Phi_f(u,\xi) \in L^2(\mathbb{R}^2)$ of a unitary energy signal $f \in L^2(\mathbb{R})$ as follows

$$\mathrm{H}_\alpha(\Phi_f) = \frac{1}{1-\alpha}\ \log_2 \iint \Phi_f^\alpha(u,\xi)\mathrm{d}u\mathrm{d}\xi\ . \tag{4}$$

The application to our problem is related to the concept that minimizing the complexity or information of a set of time-frequency representations of a same signal is equivalent to maximizing the concentration, peakiness, and therefore the sparsity of the analysis.

3.1 Entropy Evaluation on Spectrograms with Raised Cosine Windows

The method of minimizing the entropy of a time-frequency density to achieve sparsity has shown to give interesting results both analytically and numerically: in [11] some results are demonstrated with signal composed by Gaussian atoms and a Gaussian window for the spectrogram analysis. We have obtained further results with impulses, sinusoids with constant frequency and with linearly varying frequency, using the class of the so called *raised cosine* analysis windows, which is a common choice in audio applications: they guarantee good features of time-frequency localization and a good tradeoff between the main and the side lobes in their Fourier transform (for a list of features characterizing the window functions see [14] chapter IV). We have focused in particular on the *Hanning window*

$$h(t) = \cos^2(\pi t)\chi_{[-\frac{1}{2},\frac{1}{2}]} \tag{5}$$

with χ the indicator function of the specified interval, but it is possible to generalize what we demonstrate to the entire class considered.

We consider different scaled versions of a Hanning window function

$$h_{\lrcorner l}(t) = \frac{1}{\sqrt{l}} \cos^2\left(\frac{\pi t}{l}\right)\chi_{[-\frac{1}{2l},\frac{1}{2l}]} \tag{6}$$

with $l \in L$, a finite set of positive real values. Here we analyze the case of a sinusoid with constant frequency ξ_0

$$s(t) = e^{i2\pi\xi_0 t} \ . \tag{7}$$

The spectrogram of s cannot be defined as in (2) because s is not a function in $L^2(\mathbb{R})$; we preserve the same formula assuming that signals have finite duration, so considering that s is zero outside of a finite interval. For each one of the scaled windows we obtain

$$\mathrm{PS}_{sl}(u,\xi) = \left| \int s(t)h_{\lrcorner l}(t-u)e^{-2\pi i\xi t}\mathrm{d}t \right|^2 = \frac{l}{4}\left|\frac{\mathrm{sinc}(\pi l(\xi-\xi_0))}{1-l^2(\xi-\xi_0)^2}\right|^2 \ . \tag{8}$$

As above, here is not possible to apply the definition (4) as the time integral would not converge; but the densities in (8) are not time dipendent, so we can define an *instantaneous* entropy as

$$\mathrm{H}^*_\alpha(\Phi_f) = \frac{1}{1-\alpha}\ \log_2 \int \mathrm{PS}^\alpha_{sl}(u,\xi)\mathrm{d}\xi \ , \tag{9}$$

obtaining the general entropy with a second integration on finite time intervals. For every spectrogram obtained in (8) we have

$$\mathrm{H}^*_\alpha(\mathrm{PS}_{sl}) = \frac{1}{1-\alpha}\left(\log_2 \int \left|\frac{\mathrm{sinc}(\pi\xi')}{1-\xi'^2}\right|^{2\alpha}\mathrm{d}\xi' + \log_2\left(\frac{l^{\alpha-1}}{2^{2\alpha}}\right)\right) \ , \tag{10}$$

where $\xi' = l(\xi-\xi_0)$. We see that the first term does not depend on l; if we fix α, the second term is decreasing with $l \in \mathbb{R}^+$.

The sparsity measure we consider is obtained minimizing the entropy of a set of time-frequency densities, so in this case we are interested in finding l^* such that

$$l^* = \min_{l\in L}\ \mathrm{H}^*_\alpha(\mathrm{PS}_{sl}) \ , \tag{11}$$

which for every given order α is $l^* = \max_{l \in L} l$. This means that for a sinusoid with constant frequency ξ_0 and for any order α the method gives as optimal spectrogram the one taken with the larger scaling l in the finite set specified. As such a signal is perfectly concentrated in the frequency domain, this choice corresponds to our needs since it provides the best frequency resolution available. With a similar procedure we have analytically and numerically investigated the case of an impulse $s(t) = \delta(t - t_0)$: the method identifies as optimal spectrogram the one obtained with the smaller scaling, so providing best time resolution; as above, the choice of the optimal window is not dependent on the order α.

For a sinusoid with linearly varying frequency $s(t) = e^{i2\pi(\xi_0 t + at^2)}$ the optimal spectrogram is chosen according to the slope a: for higher slopes time precision is enhanced, while frequency resolution is maximized for lower. The optimal analyses obtained for different classes of signals with different entropy orders have numerically confirmed that a dependency of the method on α is shown as soon as the signal becomes more complex. The value $\alpha = 0.7$ has been heuristically fixed for the example in figure 2.

4 Frame Theory

The spectrogram (2) is defined as a function in $L^2(\mathbb{R}^2)$ and it can be interpreted as an energy density; Frame Theory ([2],[10]) extends the concept of orthonormal basis in a Hilbert space, and in our domain it provides a theory for the discretization of time-frequency densities (see [14],[6]).

4.1 Basic Definitions

Given a Hilbert space H seen as a vector space on $\mathbb{C}$, with its own scalar product, we consider in H a set of vectors $\{\phi_\gamma\}_{\gamma \in \Gamma}$ where the index set Γ may be infinite and γ can also be a multi-index. The set $\{\phi_\gamma\}_{\gamma \in \Gamma}$ is a *frame* for H if there exist two positive non zero constants A and B, called *frame bounds*, such that for all $f \in H$,

$$A\|f\|^2 \leq \sum_{\gamma \in \Gamma} |\langle f, \phi_\gamma \rangle|^2 \leq B\|f\|^2 \ . \tag{12}$$

We are interested in the case $H = L^2(\mathbb{R})$ and Γ countable, as it represents the standard situation where a signal f is decomposed through a countable set of given functions $\{\phi_k\}_{k \in \mathbb{Z}}$. The frame bounds A and B are the infimum and supremum, respectively, of the eigenvalues of the *frame operator* U, defined as

$$\mathrm{U}f = \sum_{k \in \mathbb{Z}} \langle f, \phi_k \rangle \phi_k \ . \tag{13}$$

A *Gabor frame* is obtained by time-shifting and frequency-transposing a window function g according to a regular grid. Given a time step a and a frequency step

b we write $\{u_n\}_{n\in\mathbb{Z}} = an$ and $\{\xi_k\}_{k\in\mathbb{Z}} = bk$; these two sequences generate the nodes of the time-frequency grid for the frame $\{g_{n,k}\}_{(n,k)\in\mathbb{Z}^2}$ defined as

$$g_{n,k}(t) = g(t-u_n)e^{2\pi i\xi_k t} \; ; \tag{14}$$

the nodes are the centers of the Heisenberg boxes associated to the windows in the frame. The grid has to satisfy certain conditions ([5]) for $\{g_{n,k}\}$ to be a frame, which impose limits on the choice of the time and frequency steps. There is clearly a relation between the steps a, b and the frame bounds A, B, so that the frame bounds provide an information on the redundancy of the decomposition of the signal within the frame.

Frames allow to decompose a signal using several different classes of functions: the choice of a specific frame shapes the information that we can earn about the signal through its decomposition. For any frame $\{\phi_k\}_{k\in\mathbb{Z}}$ there exist dual frames $\{\tilde{\phi}_k\}_{k\in\mathbb{Z}}$ such that for all $f \in L^2(\mathbb{R})$

$$f = \sum_{k\in\mathbb{Z}} \langle f, \phi_k\rangle \tilde{\phi}_k = \sum_{k\in\mathbb{Z}} \langle f, \tilde{\phi}_k\rangle \phi_k \; , \tag{15}$$

so that given a frame it is always possible to perfectly reconstruct a signal f using the coefficients of its decomposition through the frame. For a Gabor frame built with a window function g, such coefficients are given by sampling the STFT of f with window g according to the nodes of the time-frequency grid of the frame.

4.2 Multiple Gabor Frames

In our adaptive framework, we look for a method to achieve an analysis with multiple resolutions: thus we need to combine the informations coming from the decompositions of a signal in several frames of different window functions. *Multiple Gabor frames* have been introduced in [17] to provide the original Gabor analysis with flexible multi-resolutions techniques: given a set of index $L \subseteq \mathbb{Z}$ and different frames $\{g^l_{n,k}\}_{(n,k)\in\mathbb{Z}^2}$ with $l \in L$, a multiple Gabor frame is obtained with a union of the single given frames. The different g^l do not necessarily share the same type or shape: in our method an original window is modified with a finite number of scaling as described in (6); then all the scaled versions are used to build $|L|$ different frames which constitute the initial multi-frame.

A Gabor multi-frame has in general a significative redundancy which lowers the readability of the analysis. This limit has been overcome with the definition of *reduced multi-frames* ([6]), which turn out to belong to the recently introduced class of *quilted frames* ([7]): a decomposing system is obtained with an union of certain subsets of the individual frames in a multi-frame. The choice of the subsets is realized to locally privilege specific resolutions, and in order to assure that the resultant system is still a frame. Such a variable structure introduces some difficulties in the analytic determination of a dual frame and a reconstruction formula as in (15). The system we use for the decomposition process is actually a quilted frame, whose local structure is determined automatically by the sparsity evaluations. On the other hand, our reconstruction method is not defined by deducing a dual frame: we extend the technique introduced in [9] instead, as detailed in the next section.

5 An Algorithm for Spectrogram Time-adaptation

We now summarize the main operations of our algorithm providing an example of time-adaptive spectrogram performed on a B4 note played by a marimba.

Given a finite set L of scaling values, we create a Gabor multi-frame with $|L|$ different scaled versions of a Hanning window. The choice of the time-frequency steps for the individual frames determines the structure of the energy densities whose sparsities have to be evaluated: the spectrograms obtained with the different windows are sampled according to the grids thus obtained. We perform a local entropy calculation on every discretized spectrogram: taken a certain time-frequency rectangular subset of the whole grid, we calculate the entropy of the spectrogram coefficients in the subset, normalized to obtain a density with values between 0 and 1. We have numerically verified that as long as the decomposing system is a frame the entropy measure is invariant to redundancy variation: chosen a frame originated by a certain window function, the entropy of the spectrogram of a signal does not increase if the frame grid is tightened. Therefore, a choice for a practical implementation is to consider the same time-frequency steps a and b for all the systems: to guarantee that all the $|L|$ scaled windows constitute a frame when translated and modulated according to this global grid, the time step a must be set with the hop size assigned to the smallest window frame. On the other hand, as the FFT of a discrete signal has its same number of points, the frequency step b has to be the FFT size of the largest window analysis: for the smaller ones, a zero-padding is performed.

Another remark concerns the local entropy evaluation: if we choose a rectangular subset of the whole grid, the coefficients of the spectrograms there contained do not correspond to the same part of signal, as windows have different time supports. Therefore, a preliminary weighting of the signal has to be performed before the calculations of the local spectrograms: this step is necessary to balance the influence between coefficients which regard parts of signal shared or not shared by the different frames analyses.

Adaptation is then obtained along the time dimension by evaluating and minimizing the entropy over such rectangular subsets, which cover the whole frequency grid: they are chosen with a fixed time step and in order to have a constant overlap between each two consecutive subsets in the series.
For every signal segment individuated by the rectangular subsets of the time-

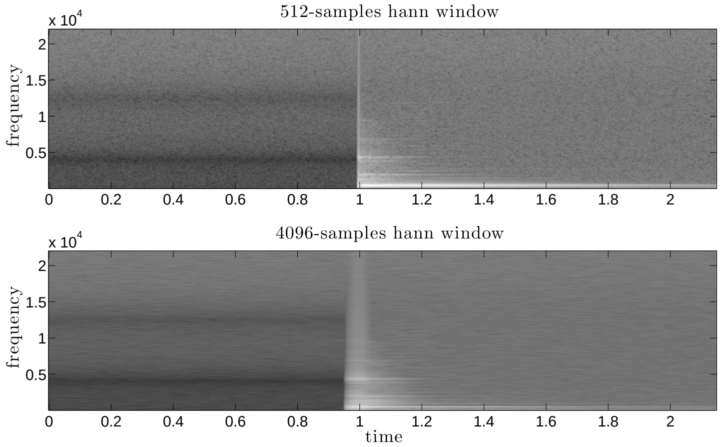

Fig. 1. Two different spectrograms of a B4 note played by a marimba, with Hanning windows of sizes 512 (top) and 4096 (bottom) samples.

frequency grid, the sparsest local analysis is defined to be the one with minimum Rényi entropy: the best window is thus defined consequently. The global signal time adapted analysis is realized opportunely assembling the slices of local sparsest analyses: they are obtained with a further spectrogram calculations on the unweighted signal employing the chosen best window.

In figure 1 we give an example of an adaptive analysis performed by our algorithm with four Hanning windows of different sizes on a real instrumental sound, a B4 note played by a marimba: this sound combines the need for a good time resolution at the moment of the percussion, with that of a good frequency resolution on the harmonic resonance of the instrument. This is fully provided by the algorithm, as shown in the adaptive spectrogram at the bottom of the figure 2.

To complete the description of our framework we finally describe the technique that we use for the reconstruction of the signal through the coefficients of its adapted analysis. The re-synthesis method introduced in [9] gives a perfect reconstruction of the signal as a weighted expansion of the coefficients of its STFT in the original analysis frame. Let $S_f[n,k]$ be the sampled STFT of a signal f according to a Gabor frame as in (14), with window function g and time step a; fixing n, an iFFT expansion gives a windowed segment of f

$$f_g(n,l) = g(na-l)f(l) \ , \tag{16}$$

whose time location depends on n. An immediate perfect reconstruction of f is given by

$$f(l) = \frac{\sum_{n=-\infty}^{+\infty} g(na-l) f_g(n,l)}{\sum_{n=-\infty}^{+\infty} g^2(na-l)} \ . \tag{17}$$

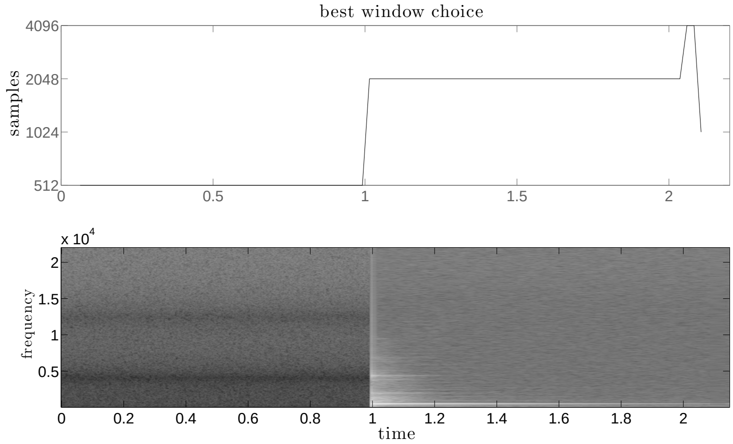

Fig. 2. Example of an adaptive analysis performed by our algorithm with four Hanning windows of different sizes (512, 1024, 2048 and 4096 samples) on a B4 note played by a marimba: on top, the best window chosen as a function of time; at the bottom, the adaptive spectrogram with $\alpha = 0.7$.

We extend the same technique using a variable window g according to the composition of the reduced multi-frame, obtaining a perfect reconstruction as well. The interest of (17) is that the given distribution needs not to be the STFT of a signal: for example, a transformation $S^*[n, k]$ of the STFT of a signal could be considered. In this case, (17) gives the signal whose STFT has minimal least squares error with $S^*[n, k]$.

6 Conclusions

We have shown how an adaptive analysis/synthesis framework can be defined combining two main models: Frame Theory to rule the time segmentation in the analysis process, Rényi entropies to define a class of sparsity measures. The choice of a common time-frequency grid for all the frames allows a practical implementation and let the different spectrograms contain the same number of coefficients. Nevertheless this requires a substantial pre-weighting of the signal and a high overlap for larger windows which has a significative computational cost. For this reason we are interested in defining strategies for a specific choice of the time-frequency grid related to every different frame involved in the analysis.
A further fundamental improvement will concern the extension of our method to perform automatic adaptation of the resolution simultaneously in time and frequency: this will require the conception of new re-synthesis techniques managing time-frequency overlap between different frames.

References

1. Baraniuk, R.G., Flandrin, P., Janssen, A.J.E.M., Michel, O.J.J.: Measuring Time-Frequency Information Content Using the Rényi Entropies. IEEE Trans. Info. Theory vol.47, n.4, (2001)
2. Christensen, O.: An Introduction To Frames And Riesz Bases. Birkhäuser, Boston, (2003)
3. Cohen, L.: Time-Frequency Distributions-A Review. Proceedings of the IEEE vol.77, n.7, (1989)
4. Cohen, L.: Time-Frequency Analysis. Prentice-Hall, Upper Saddle River, (1995)
5. Daubechies, I.: The Wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE Trans. Info. Theory vol.36, n.5, (1990)
6. Dörfler, M.: Gabor Analysis for a Class of Signals called Music. PhD thesis, NuHAG, University of Vienna, (2002)
7. Dörfler, M.: Quilted Gabor frames - a new concept for adaptive time-frequency representation. eprint arXiv:0912.2363, (2010)
8. Flandrin, P.: Time-Frequency/ Time-Scale Analysis. Academic Press, San Diego, (1999)
9. Griffin, D.W., Lim, J.S.: Signal Estimation from Modified Short-Time Fourier Transform. IEEE Trans. Acoust. Speech Signal Process vol.32, n.2, (1984)
10. Gröchenig, K.: Foundations of Time-Frequency Analysis. Birkhäuser Boston, (2001)
11. Jaillet, F.: Représentation et traitement temps-fréquence des signaux audionumériques pour des applications de design sonore. PhD thesis, Université de la Méditerranée - Aix-Marseille II, (2005)
12. Jaillet, F., Balazs, P., Dörfler, M.: Nonstationary Gabor Frames. INRIA a CCSD electronic archive server based on P.A.O.L [http://hal.inria.fr/oai/oai.php], (2009)
13. Lukin, A., Todd, J.: Adaptive Time-Frequency Resolution for Analysis and Processing of Audio. Audio Engineering Society Convention Paper, (2006)
14. Mallat, S.: A wavelet tour on signal processing. Academic Press, San Diego, (1999)
15. Rudoy, D., Prabahan, B., Wolfe, P.: Superposition frames for adaptive time-frequency analysis and fast reconstruction. eprint arXiv:0906.5202v1, (2009)
16. Wolfe, P. J., Godsill, S. J., Dörfler, M.: Multi-Gabor Dictionaries for Audio Time-Frequency Analysis. Proc. IEEE WASPAA, (2001)
17. Zibulski, M., Zeevi, Y. Y.: Analysis of multiwindow Gabor-type schemes by frame methods. Appl. Comput. Harmon. Anal. vol.4, n.2, (1997)