JJCAAS 2009
LMA, Marseille
25, 26, 27 novembre 2009
Les 5e Journées Jeunes Chercheurs en Audition, Acoustique musicale et Signal audio
Dubium sapientiae initium. (R. Descartes)

Different resolutions for different informations

Marco Liuni*,**, Axel Röbel*, Marco Romito**, Xavier Rodet*

*IRCAM - CNRS STMS, Analysis/Synthesis Team

**Università di Firenze, Dip. di Matematica "U. Dini"

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

INTRODUCTION

The existing algorithms for **analysis, synthesis and transformation of sounds** ask the users to have a high competence on signal processing. To make the access to these efficient algorithms easier, it is desirable to **let the algorithms adapt themselves** to the signal and to the transformation needed.

One of the key parameter for nearly all the algorithms which analyze or process the sounds by using time-frequency representations or sinusoidal models is **the size of the analysis window**: here we show some examples about how this parameter is linked to the resolution of a signal analysis, and which informations can be highlighted choosing a certain size instead of another.

We are at present studying and developing algorithms for the **automatic adaptation of the analysis window size**. The first method approached is the parallel evaluation of multiple choices of the parameters and the subsequent combination of the final outcome through a comparison between their features.

ANALYSIS-SYNTHESIS SCHEME

An **audio signal** can be represented as a function

$$s : \mathbb{R} \to \mathbb{C},\ s \in L^2(\mathbb{R})\ .$$

A system for the analysis and synthesis of a signal is composed of

- a set of functions called **time-frequency atoms**
$$\{\phi_\gamma\}_{\gamma\in\Gamma} \subset L^2(\mathbb{R}) \tag{1}$$
- a **time-frequency transform** operator associated to the above set
$$T_s(\gamma) = \int s(t)\overline{\phi_\gamma(t)}\mathrm{d}t = \langle s, \phi_\gamma\rangle \tag{2}$$
- a **reconstruction formula**
$$s(t) = \int_{\gamma\in\Gamma} T_s(\gamma)\phi_\gamma(t)\mathrm{d}\gamma \tag{3}$$

FRAMES THEORY

For the reconstruction formula to exist, the set $\{\phi_\gamma\}_{\gamma\in\Gamma}$ has to satisfy a condition obtained by relaxing the definition of an orthonormal system in a Hilbert space.

Definition 1 *The set $\{\phi_n\}_{n\in\Gamma}$ is a frame of $L^2(\mathbb{R})$ if A and B exist such that they are positive and for every $s \in L^2(\mathbb{R})$*

$$A\|s\|^2 \le \sum_{n\in\Gamma} |\langle s, \phi_n\rangle|^2 \le B\|s\|^2\ , \tag{4}$$

where thanks to the features of the Hilbert space $L^2(\mathbb{R})$ we can suppose Γ to be countable.

WINDOW FUNCTIONS

We consider a **window function** as a function $h \in L^2(\mathbb{R})$ such that $\|h\| = 1$ and $h(t) = h(-t)$;
modifying h with a time translation of a factor t_0 and a frequency scaling by a factor ν_0 we obtain the set $\{h_{t_0,\nu_0}\}_{(t_0,\nu_0)\in\mathbb{R}^2}$ where

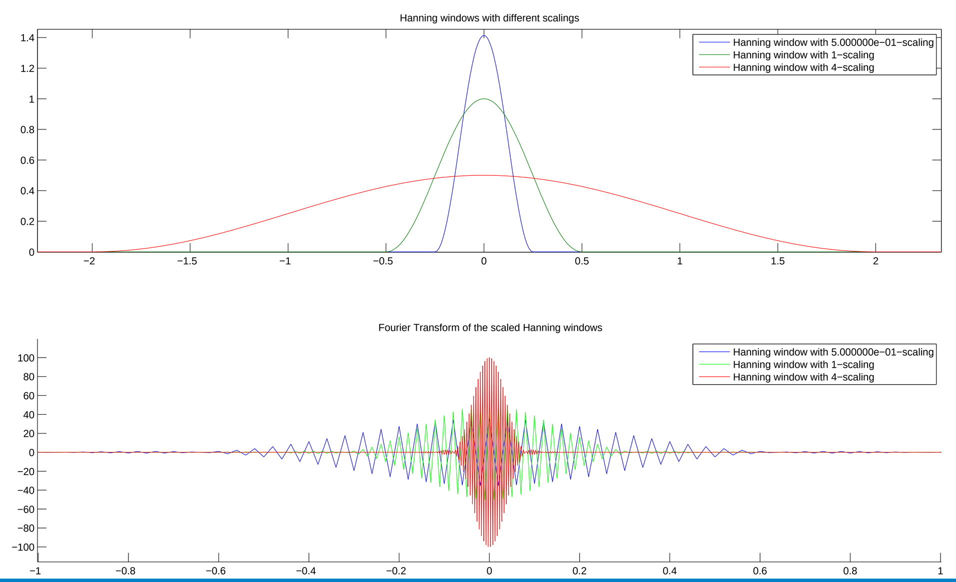
$$h_{t_0,\nu_0}(t) = h(t - t_0)e^{2\pi i\nu_0 t}\ . \tag{5}$$

Remark 1

- The **Short Time Fourier Transform** is the time-frequency transform associated to the set $\{h_{t_0,\nu_0}\}_{(t_0,\nu_0)\in\mathbb{R}^2} \subset L^2(\mathbb{R})$
- Such set contains infinite countable frames of window functions for $L^2(\mathbb{R})$

A frame can be built by conveniently combining different sets of atoms derived from different window functions.
We are particularly interested to the case of the scaling of a same window function, which modifies its **time-frequency spread**,

$$h_l(t) = \frac{1}{\sqrt{l}}h\left(\frac{t}{l}\right)\ . \tag{6}$$



ADAPTIVE SPECTROGRAM

The **Spectrogram** of a signal s is the squared modulus of the STFT of the signal

$$PS_s(t_0, \nu_0) = \left|\int s(t)h(t - t_0)e^{-2i\pi\nu_0 t}\mathrm{d}t\right|^2\ . \tag{7}$$

It is always positive and with certain exceptions it can be treated as a probability distribution.

It is possible to perfectly represent the signal s in infinite many different ways, by choosing an appropriate set of multi-index $\Gamma \subset \mathbb{R}^2$ so that the relative frame $\{h_{t_0,\nu_0}\}_{(t_0,\nu_0)\in\Gamma}$ is dense enough in the time-frequency plane.

Given PS_s^h the spectrogram of a signal s through the frame built on the window h, a **sparsity measure** can be defined to evaluate the energy concentration in the analysis PS_s^h.

The notion of **best energy concentration** is not uniquely defined and is linked to the desired trade-off between time and frequency resolution.

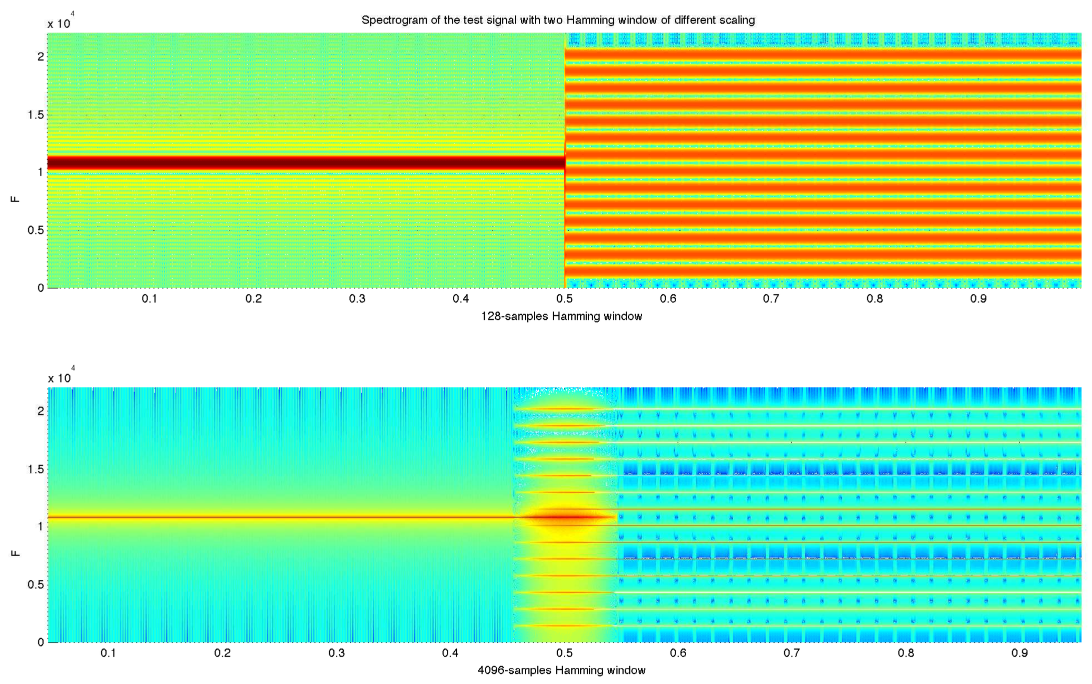


Fig.2 Spectrogram of a test signal with two Hamming window of different scaling.

As shown in [BF01] it is possible to extend the notion of **entropy** to certain time-frequency representations, such as the spectrogram, to define on them a class of sparsity measures.

Definition 2 *Given a time-frequency representation $C_s \in L^2(\mathbb{R}^2)$ of the signal s, its* Rényi entropy *of order $\alpha > 0, \alpha \ne 1$ is*

$$H_\alpha(C_s) = \frac{1}{1-\alpha}\log_2 \iint C_S^\alpha(t,\nu)\mathrm{d}t\mathrm{d}\nu \tag{8}$$

Given a signal s and n different spectrograms $PS_s^{h_1}, ..., PS_s^{h_n}$ of the same signal based on different windows $h_1, ..., h_n$ we choose as the **best spectrogram** the one which minimizes the Rényi entropy measure,

$$l^* = \arg\min_{l=1,...,n} H(PS_s^{h_l})\ . \tag{9}$$

ALGORITHM AND EXAMPLES

- Input and main parameters: the signal s, the set of window functions $h_1, ..., h_n$, the Rényi entropy order α
- Computation of the n different spectrograms
- Evaluation of the Rényi entropy values on the global analysis
- Choice of the best spectrogram

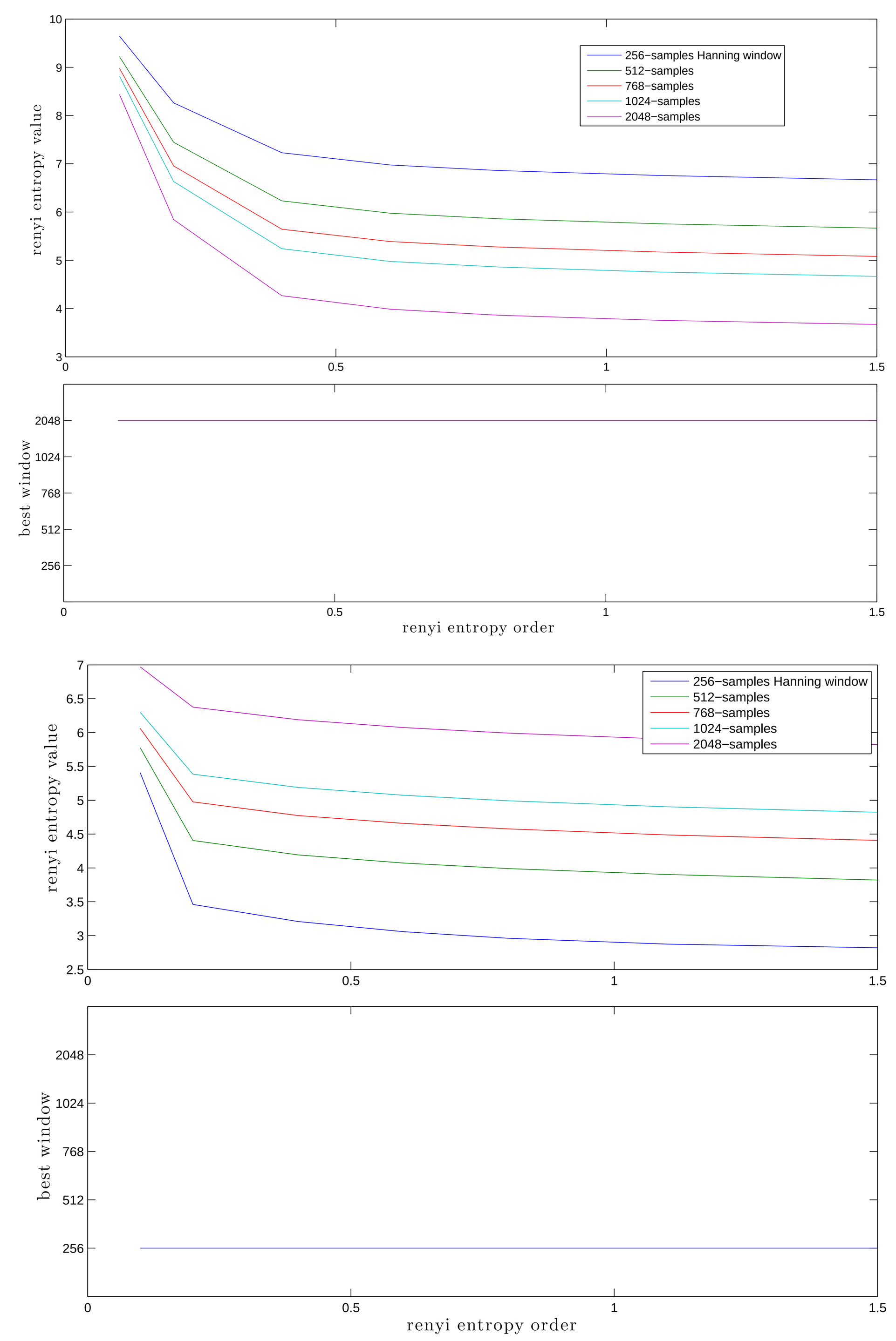


Fig.3 The choice made by the algorithm for two basic signals: two different sinusoids in the upper figure, an impulse in the lower. In these cases, the choice is α-independent

RESEARCH PROGRAM

- Development and investigation on the best strategy to set the hop size of the different spectrograms, in order to enhance the compatibility of the different analysis
- Definition of a strategy for time and frequency localization of the Rényi entropy evaluation, to obtain a local adaptation of the window size to the signal
- Investigation on the relation between the Rényi entropy order α and the choice of the best spectrogram by the algorithm

REFERENCES

- [BF01] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, O.J.J. Michel, *Measuring Time-Frequency Informaion Content Using the Rényi Entropies*; IEEE Transaction on Information Theory, vol.47, n.4, 2001.
- [Co95] Leon Cohen, *Time-Frequency Analysis*; Prentice-Hall, 1995.
- [Fl99] Patrick Flandrin, *Time-Frequency/ Time-Scale Analysis*; Academic Press, 1999.
- [Mal99] Stéphane Mallat, *A wavelet tour on signal processing*; Academic Press, 1999.