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

Marco Liuni[1,2,*], Axel Röbel[2], Marco Romito[1], and Xavier Rodet[2]

[1] Universitá di Firenze, Dip. di Matematica "U. Dini"
Viale Morgagni, 67/a - 50134 Florence - ITALY
[2] IRCAM - CNRS STMS, Analysis/Synthesis Team 1,
Place Igor-Stravinsky - 75004 Paris - FRANCE
{marco.liuni,axel.roebel,xavier.rodet}@ircam.fr
marco.romito@math.unifi.it
http://www.ircam.fr/anasyn.html

Abstract. We propose a method for automatic local time-adaptation of the spectrogram of audio signals: it is based on the decomposition of a signal within a Gabor multi-frame through the STFT operator. The sparsity of the analysis in every individual frame of the multi-frame is evaluated through the Rényi entropy measures: the best local resolution is determined minimizing the entropy values. The overall spectrogram of the signal we obtain thus provides local optimal resolution adaptively evolving over time. We give examples of the performance of our algorithm with an instrumental sound and a synthetic one, showing the improvement in spectrogram displaying obtained with an automatic adaptation of the resolution. The analysis operator is invertible, thus leading to a perfect reconstruction of the original signal through the analysis coefficients.

Keywords: adaptive spectrogram, sound representation, sound analysis, sound synthesis, Rényi entropy, 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 among others. The most challenging class of signals to consider is indeed music: the completely new perspective opened by contemporary music, assigning a fundamental role to concepts as noise and timbre, gives musical potential to every 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. Traditional methods based on single sets of atomic functions have important limits: a Gabor frame imposes a fixed resolution over all the time-frequency plane, while a wavelet frame gives a strictly determined variation of the resolution: moreover, the user is frequently

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

S. Ystad et al. (Eds.): CMMR 2010, LNCS 6684, pp. 60–75, 2011.

asked to define himself the analysis window features, which in general is not a simple task even for experienced users. 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* [22] of the spectrogram; we then define a re-synthesis technique with an extension of the method proposed in [11]. Our approach can be presented schematically in three parts:

1. a model for signal analysis exploiting concepts of Harmonic Analysis, and Frame Theory in particular: it is a generally highly redundant decomposing system belonging to the class of multiple Gabor frames [8],[14];
2. a sparsity measure defined on time-frequency localized subsets of the analysis coefficients, in order to determine local optimal concentration;
3. a reduced representation obtained from the original analysis using the information about optimal concentration, and a synthesis method through an expansion in the reduced system obtained.

We have realized a first implementation of this scheme in two different versions: for both of them a sparsity measure is applied on subsets of analysis coefficients covering the whole frequency dimension, thus defining a time-adapted analysis of the signal. The main difference between the two concerns the first part of the model, that is the single frames composing the multiple Gabor frame. This is a key point as the first and third part of the scheme are strictly linked: the frame used for re-synthesis is a reduction of the original multi-frame, so the entire model depends on how the analysis multi-frame is designed. The section *Frame Theory in Sound Analysis and Synthesis* treats this part of our research in more details.

The second point of the scheme is related to the measure applied on the coefficients of the analysis within the multi-frame to determine local best resolutions. We consider measures borrowed from Information Theory and Probability Theory according to the interpretation of the analysis within a frame as a probability density [4]: our model is based on a class of entropy measures known as *Rényi entropies* which extend the classical Shannon entropy. The fundamental idea is that minimizing the complexity or information over a set of time-frequency representations of the same signal is equivalent to maximizing the concentration and peakiness of the analysis, thus selecting the best resolution tradeoff [1]: in the section *Rényi Entropy of Spectrograms* we describe how a sparsity measure can consequently be defined through an information measure. Finally,