

Proposition for post-doctoral research (contract researcher)

Title : **Non-stationary and non-redundant time-frequency transforms for audio coding**

Laboratory : Laboratory of Mechanics and Acoustics (LMA), CNRS UPR 7051, Marseille, France.

Duration : 12 months.

Starting date: June 2014.

Funding : ANR POTION, French-Austrian collaboration.

Supervisor : Olivier DERRIEN – derrien@lma.cnrs-mrs.fr

Profile : The hired contract researcher will be a PHD in the field of applied mathematics or signal processing, with a background in time-frequency analysis. Knowledge in audio signal processing will be appreciated, but is not mandatory.

Detailed specification

1. The ANR POTION project

POTION is a research project funded by the French National Agency for Research (ANR) and the Austrian Science Fund (FWF), which deals with perceptual optimization of time-frequency representations and application to compressive audio coding (like MP3). The scientific program is divided in three main tasks: Studying new non-stationary time-frequency representations, building new time-frequency auditory masking models and finally applying this knowledge to audio coding. The proposed study is part of the first task, and will be conducted both at the Laboratory of Mechanics and Acoustics (LMA) in Marseille, France and the Acoustics Research Institute (ARI) in Vienna, Austria. Several scientific workshops in Vienna will be organized during these 12 months.

2. Non-stationary time-frequency transforms

Time-frequency representations (or transforms) are widely used in (non-stationary) signal processing. They allow to localize the signal energy in the time-frequency plane. The most popular time-frequency representation is the “Short-Time Fourier Transform” which sampled version in time and frequency is called “Gabor Transform”. The Gabor transform has many advantages: it has perfect reconstruction capability, and there is a large choice for analysis and synthesis windows, which allows to control the energy spreading in the time-frequency plane. Recently, non-stationary versions (in time and in frequency) of the Gabor transform have been proposed by the Austrian partner of the project. It is now possible to decompose the signal in the time-frequency domain with non-linear scales in time or in frequency. This is especially interesting for audio coding, because it allows the frequency scale to mimic the auditory spectral analysis (for instance the Bark or ERB scale). Unfortunately, these transforms introduce redundancy, which means that the density of transform coefficients is much higher than the density of samples in the original time-domain.

3. The issue of redundancy and application to compressive audio coding

The goal of compressive audio coding is to reduce the amount of binary information used to store or transmit the audio signal. This is performed by a quantizer applied on the transform coefficients. Thus, the transform must not introduce redundancy. Otherwise, it would not be possible to compensate for this disadvantage in the quantization stage. Most audio coders rely on transforms that belong to the family of Discrete Cosine Transforms (DCT). These do not introduce redundancy because they are equivalent to a change of basis. Thus, each non-redundant transform is associated to a time-frequency basis (where each vector has a well-localized energy in both time and frequency). Bases associated to DCTs are called “Malvar bases”. At present, time-frequency bases all have linear frequency scales.

4. The goal of the study

In the first task of the POTION project, one goal is to study the possibility to define time-frequency bases (or equivalently non-redundant time-frequency transforms) with non-linear time and frequency scales. We believe that this could be achieved by modifying classical bases from the literature (e.g. Wilson bases or Malvar bases), using methods that are similar to the ones used for building non-stationary Gabor transforms. We also aim at characterizing such bases and relate them to the non-stationary Gabor transform.