

Adapted representations of audio signals for music instrument recognition

Pierre LEVEAU

My current research deals with the parameterization of audio signals for music instrument recognition in real playing conditions (not isolated notes). This work is developed at the Laboratoire d'Acoustique Musicale (Paris) with Laurent Daudet, in narrow collaboration with Gaël Richard from ENST (Paris). The first lab brings its experience in acoustics and music perception, the second one its expertise in audio signal processing, especially in information retrieval. It constitutes the final project of a DEA (equivalent to a research oriented master) prepared at IRCAM (Paris), and must be followed by a Ph.D.

We will first focus on an improvement of the existing parameterizations, by distinguishing the transients and release parts of music signals. Then a new way to parameterize audio signals will be explored: the idea is to develop heavily redundant representation spaces, in which the representation of the signal would be parsimonious. For instance, this space can combine local Fourier transform and Wavelet coefficients.

The first step of this work is to segment audio signals into notes. It means that each segment will contain only one transient, followed by the sustained part of the note. This goal is a bit ambitious for the polyphonic signals, but the results are expected to be statistically exploitable. To achieve this objective, using detection functions has proved to be the most convenient, being the commonly used method to detect musical onsets.

As a first approach, we will consider solo phrases. A few different detection functions have been tested on a reference music signals database, with a reference hand-labeled onset list for each file. The hand-labeling process has been precisely studied [article submitted at ISMIR 04] because a detection function evaluation is heavily dependent on the reference put by hand.

Once the appropriate detection function is found, the music instrument recognition algorithms are tested only on the detected transients of the signals. The results will give the adapted parameters to represent these transients. The same operations can be practiced only on the sustained parts, and will also give the adapted parameters. Once all the consistent parameters have been chosen for these two sets of segments, music instrument recognition can be performed on the combined sets of features.

The following work will consist in finding appropriate algorithms to find the representation in the redundant space. The algorithms will also be based on segmented notes. In the general case, commonly used algorithms like matching pursuit give good results but their computation is very costly. Adapted algorithms for audio signals, particularly music notes, have also to be developed.