MODELING THE SPECTRUM STRUCTURE WITHIN THE NMF FRAMEWORK. APPLICATION TO INHARMONIC SOUNDS

Bertrand David[1], *Francois Rigaud*[2], *Laurent Daudet*[3]
[1] Institut Mines-Télécom; Télécom ParisTech; CNRS LTCI , France
[2] Audionamix, France
[3] Institut Langevin; Paris Diderot Univ.; ESPCI ParisTech; CNRS, France
bertrand.david@telecom-paristech.fr
francois.rigaud@audionamix.com
laurent.daudet@espci.fr

ABSTRACT

The Non negative Matrix Factorization has received much attention during the last decade, with some remarkably successful applications in the field of audio source separation or automatic music transcription. To that aim, prior information and modeling have often been included in the general framework to account for our knowledge on the spectral structure of sounds, such as the harmonicity, characteristic of a number of musical sounds, or the smoothness of their spectrum envelope. These help the algorithm to approach a relevant solution. The NMF indeed decomposes a non negative time-frequency representation of a musical scene into the product of low rank non negative matrices: the matrix of spectral atoms or templates and the matrix of their time activation. But this decomposition is not unique and moreover it is not rare that the convergence reaches a local minimum of the cost function where the spectral atoms are not easily identifiable, or at least, have to be post-processed either to separate the contributions of different audio events or to aggregate them to recover a single, coherent, musical note. This presentation will review techniques we developed in the past few years to model the spectrum structure within the NMF framework and their application to analyze inharmonic sounds, such as those of the piano. This allows us to examine some of large scale properties that characterize the state of the instrument, like the design of its tuning, or the inharmonicity curve along the whole compass.