Discriminant Approach within Non-negative Matrix Factorization for Musical Components Recognition
Zafar Rafii, Raphael Blouet & Antoine Liutkus
Mist Technologies R&D - Paris Cyberville - 204 rue de Cimède, 75019 Paris, France
firstname.lastname@mist-technologies.com

Application context
- Mist Technologies develops systems for single-source separation - Upmix for professionals (DVD reissues, Movies: la Vie en Rose) and general public applications (www.songcooker.com).
- Prior information on mixture components allows to obtain good performance but requires manual instrument identification. This task is handmade and very costly.
- As an instrument detection in polyphonic recordings appears to be a very difficult task, we first focus on close-set instruments identification. Given a close set of musical instruments, it is possible to automatically identify some components in polyphonic music.
- This system could also be used as a general front-end to any high-level information retrieval system (gene detection, transcription...)

Scientific context
- Investigation of the use of Non-negative Matrix Factorization (NMF) to model audio signals.
- Virtanen takes advantage of NMF for sound source separation [Virt06].
- Smaragdis proposes a modified version of the NMF algorithm which is able to identify components with temporal structure [Sma04].
- Cost (CD07) applies NMF on modulations, but also on spectrograms for pitch estimation and instrument recognition.
- Benetos, Kotti and Kotropoulos use NMF on matrix of non-negative feature vectors extracted from audio files for instrument classification [BKSK06].

Enhancements
NMF is not unique: the factorization depends not only on the update rules but also on the starting point.

Initialization
- Random positive matrices
- Spherical K-means clustering
- Nonnegative Double Singular Value Decomposition (NNDSVD) [BG07]
- Experimental results show that NNDSVD favors components whose gains are sparse and slowly varying [Vir06].

Number of Components
There is no reliable method for the automatic estimation of the number of components. We determine it by applying a first coarse NMF, using a large number of components. To that end, Virtanen's NMF, initialized with a NNDSVD, with a strong sparseness criteria helps to keep only the components with relevant energy. These components are then filtered from their null values and used as an initialization for a second standard NMF.

Sparserness/Smoothness Constraints
Additional constraints can also be included in the update rules to enforce a convergence.

Forced Discriminant NMF
We use our enhanced NMF algorithm described above to extract the fixed magnitude spectrum components from a training database formed of K=5 instrument classes (Bass, Drums, Piano, Saxophone, Trumpet), and stored them in dictionaries.

Evaluation & Conclusions
The test set consists of 150 musical excerpts from the RWC database. Evaluation is done using 5 target instruments: Bass, Drums, Piano, Saxophone, & Trumpet.

Figure 5: Error approximation & Between-class Scatter for NMF & FDNMF

References