## Domain : Natural Language, Speech and Audio Processing

Domain extra : machine learning, signal processing
**Year :** 2010
**Starting :** automn 2010
**Status :** open

### Subject : Automatic indexing of singing voice

**Thesis advisor :** Martine Adda-Decker

Co-advisors : Jean-Luc Rouas, LIMSI/CNRS

Collaborations : CREM (Centre de Recherche d'Ethnomusicologie), LAM (Lutheries - Acoustique - Musique), Institut Jean Le Rond d'Alembert, Paris

**Abstract** :
The automatic indexing of multimedia documents on the WEB raises the status of the singing voice : should these types of acoustic segments be labelled as music, as speech or more specifically as singing voice ? If they can be automatically identified as singing voice, is it possible to automatically identify the language of the lyrics ? transcribe some of the lyrics ?
The proposed subject aims at adapting methods and techniques available for the processing of speech and music to the singing voice. Two separate investigations are proposed : the first one makes use of a corpus of singing voice from a high-quality classical repertoire, to improve the characterization of singing voice with respect to spoken voice and to propose methods for singing voice modeling and automatic language identification of the singing voice. A second investigation will make use of a huge collection of ethnomusicological recordings mixing speech and singing voice, with a large variety of distinct singing techniques, music and noise. The challenge here is to automatically separate human voices from other music or background noises and to classify human voices as either speaking or singing. The objective here is to improve the access of a unique historical, ethnomusicological collection to a large public of interested researchers.
Context (Current state of the art in the domain - max : 1000 chars) :
Automatic indexing of singing voice is an relatively new research topic.

Objectives (Scientific objectives - max : 1000 chars) :
Contribute to our knowledge of spoken/singing voice, the characterisation of different singing techniques and their improved automatic processing. The automatic indexing of speech/song/music audio documents enables new applications/usages of archives as well as of the WEB.

Work program (Main expected steps - max : 1000 chars) :
Singing voice only audio corpus :
• acoustic characterisation of sung phonemes as opposed to spoken phonemes in different languages.
• model estimation and automatic language identification.
Ethnomusicological recordings :
• voice/music/noise classification
• voice : speech/singing classification
• model estimation and automatic language identification

Extra information (Extra information like a link to a full description of the subject - max : 500 chars) :
Prerequisite (Expected background) :
**Expected funding :** research contract
Status of funding : expected
Candidates (informations on potential candidates - max : 500 chars) :

**user : martine.adda-decker@limsi.fr**