

MIREX AUDIO CHORD DETECTION (first-draft short version)

Maksim Khadkevich

FBK-Irst, Università degli studi di
Trento, Trento, Italy
khadkevich@fbk.eu

Maurizio Omologo

Fondazione Bruno Kessler - Irst
Via Sommarive, 18 - Povo - 38050
Trento, Italy omologo@itc.it

ABSTRACT

This paper describes the submission to the MIREX'10 (Music Information Retrieval eXchange) Audio Chord Detection task. Feature set, training technique, recognition schemas are briefly described. This submission is based on our submission to MIREX 2009.

1. DESCRIPTION OF THE SYSTEM

1.1. Training

For training hidden Markov models (HMM) chroma vectors are used as features. Separate models are built for each chord distinguished by the system. Each model can be characterized by a number of Gaussians in 12 dimensions, described by its mean vector and covariance matrix. Training is performed using the specific application of EM the expectation maximization (EM) algorithm – the Baum-Welch, or forward-backward algorithm. As opposed to our MIREX2009 submission, two streams of feature vectors are used in the HHMs. They correspond to 2 frequency ranges (bass and treble). To model chord transition probabilities, a 3-gram language models are trained on the given for training ground-truth material.

1.2. Recognition

Before recognition process, feature vectors from test data are extracted. The trained models are used to determine a chord labeling for each song. The Viterbi algorithm [1] is used to recognize chords from test data. It outputs a lattice which is subsequently rescored, applying language modeling weight.

2. IMPLEMENTATION AND EXPERIMENTS

Feature extraction system was implemented in Java. The Hidden Markov Model Toolkit (<http://htk.eng.cam.ac.uk/>) was used to perform training and testing. The “pre-trained” system was trained on the 180 songs of Beatles. Training of the “train-test” submission is supposed to be performed by IMIRSEL.

3. REFERENCES

- [1] L. R. Rabiner "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE. 77, 257-286.