

The FASTLab Music Analysis Kernel

The FASTLab Music Analysis Kernel (FMAK) is a software package for building and using music and sound databases. It consists of four main interfaces: analysis, segmentation, clustering, and classification. The FMAK analyzer computes both low-level and high-level features (called feature vectors or meta-data) from musical selections. The segmenter takes these feature vectors and finds the phrase, verse, and section breaks in music, thus discovering the musical form and allowing us to reduce the number of feature vectors we need to store. The clustering functions support data mining in large databases of feature vectors by grouping the data into well-defined genre clusters. The classifier adds customizable database pruning and run-time distance metrics for using genre databases. These four components can be used in a variety of ways to build software applications that process large volumes of multimedia data.

Introduction

Musical sound is a complex signal, and there are many techniques available to analyze music for the purpose of extracting high-level properties (“features”) of a musical selection. Depending on the application, any of several kinds of musical meta-data may be required, and it is the purpose of the FASTLab Music Analysis Kernel (FMAK) to perform this analysis. FMAK is a C++ software library that uses best-of-breed signal analysis, segmentation, classification, and clustering techniques for detailed music feature extraction. Applications use FMAK by making calls to the objects in the FMAK class library.

FMAK is intended for use in music and sound databases, listener preference-matching, sound finger-printing, music thumb-nailing, search engines, signal data reduction, media keying and indexing, video segmentation, and other cases where a multidimensional characterization of a music/sound selection is needed. It uses a number of different analysis techniques to extract features from musical selections.

The present version is oriented towards the analysis of popular music songs stored as uncompressed sound files, and the generation of a database of analysis feature vectors. The analysis engine is, however, style- and genre-independent, and in fact, many styles were used to test it during its development,

including popular music, contemporary art music, non-western music, and sound effects.

As the world’s entire musical inventory, both past and present, is now (or soon will be) digitized and available online in one form or another, there are more and more applications areas for the kind of perceptual feature extraction technology provided by FMAK.

The FASTLab Technology

The FASTLab Music Analysis Kernel (FMAK) is a collection of software classes for music/sound analysis, segmentation, clustering, and classification

In effect, FMAK “listens” to music and can derive over 100 attributes related to the musical structure, instrumentation, style, recording and production, and other properties.

The core analysis API uses a range of sound signal analysis techniques in the time and frequency domains, including RMS windowed envelope extraction, FFT analysis and spectral statistics, wavelet decomposition and rhythm tracking, LPC analysis and noise residual spectrum, and statistics combining these metrics.

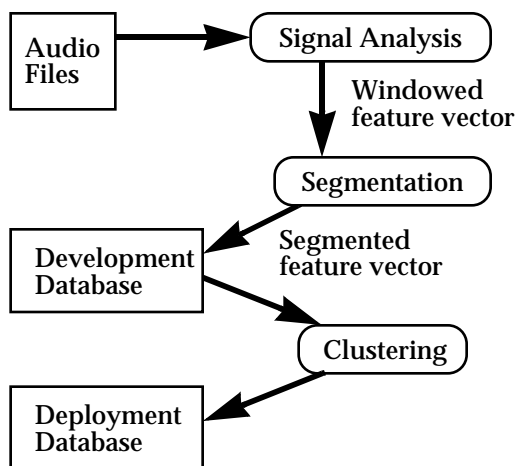
The FMAK core calculates this feature vector for many “windows” or time-slices of a song (e.g., 10 times a second), meaning that the size of the initial feature vector data set can be quite large.

Using the time-slice feature vectors, FMAK can break a musical selection into segments or sections, typically locating the rough musical structure or song verses. The segment boundaries are then used to reduce the number of feature vectors we use to characterize a song; i.e., one can store one vector for the average of the entire song, and one per verse or chorus, rather than keeping all of the (10-per-second) initial feature vectors.

For applications requiring a reference database of content, the FMAK clustering API allows developers to use standard data mining techniques to “prune” a database, grouping the data items into clusters, and then replacing the cluster with a single “representative.” The representatives may be labelled with descriptive information that can be used by a later-stage (e.g., run-time) classifier.

The Figures below show the interaction of the basic components of FMAK for a development system and a run-time application.

In the first Figure, a selection (song) is analyzed, and the (very large) set of detailed window feature vectors is passed to the segmenter. This in turn finds the section boundaries in the signal, and compacts the feature vector data that goes to the database.

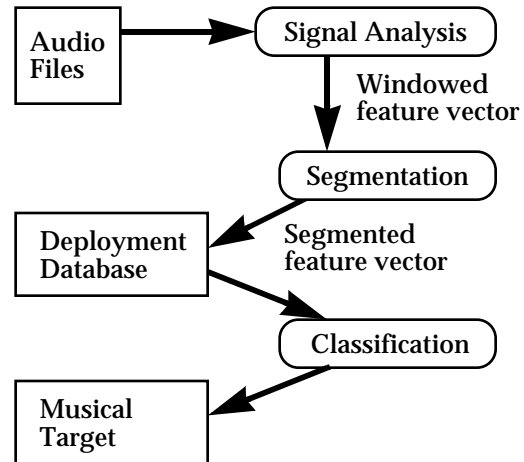


FMAK Development: Analysis, Segmentation, Clustering

In the next stage, the clusterer identifies groups of content descriptors in the database, and allows the user to label them as representing musical genres. This is used to create

a greatly reduced “deployment” database.

The second Figure shows the configuration of a run-time system that uses the reduced database with a simplified analysis/segmentation engine. The output is an item from the database that matches the given selection according to some selection criteria.



FMAK Runtime: Analysis, Classification, Selection

Summary

There are several unique features of our approach to music information retrieval:

- Fine-grained analysis;
- Sophisticated signal statistics;
- Beat, tempo, and time-domain analysis;
- Segmentation, feature data reduction;
- State-of-the-art clustering;
- Database pruning, classification, genre labeling; and
- Run-time classification.

We are very creative in using these technologies to the competitive advantage of our clients.

Contact Information

Stephen Travis Pope
 FASTLab LLC
 Tel: (805) 967-2621
 Email: stp@mat.ucsb.edu