

Report

Mattia Lecci, Federico Meson, Victor Cercos Llombart

November 7, 2017

1 Achievements and Changes

The project is following the established structure and milestones. The next step is to perform a first prototype test and weight the results. As of now we are not thinking about changing the original plan. As previously said, if we still have time after we properly finish the current single chord recognition algorithms, we might try to attack the full-song problem using the Hidden Markov Model approach.

2 Techniques

As planned we rely on *Chroma toolbox* [chr] to actualize the feature extraction. We decided to focus only on three kinds of features, specifically CENS, CLP and CRP. Given an audio file, assuming it only contains a single chord, we are able to build a (1x12) vector which describes the different chroma levels (A, A#, ..., G#). Therefore different methods for chord recognition are applied.

2.1 Template-based Chord Recognition

The first method we describe is the *template-based* method.

We have chosen two different templates to be used: the Binary and the Harmonic. In the Binary Template the different Notes are described only with 1's and 0's. In the Harmonic Template the harmonics of each Note are also taken into account. We calculate the first 6 harmonics of each notes using the Emilia Gomez [Góm06] proposed method. Following the procedure we compute the binary and the harmonic template for each chord. In order to make a decision we use the cosine distance between the extracted feature vector and the template, choosing the minimum distance chord as a prediction.

2.2 Statistical Chord Recognition

In addition to the template-based method previously described we chose to realize two different statistical models.

The first statistical model that we implemented is the *Gaussian Mixture Model*. We perform a supervised training and we build up the different Gaussian distribution. Each of these Gaussian distribution is referred to a different chord. Given a testing chord we establish the probability that it belongs to a Gaussian distribution or another computing the so-called *Mahalanobis distance*. We notice that Matlab provides methods to build the Gaussian Mixtures and also to compute the *Mahalanobis distance*. We can improve the performance of this model using more training samples and therefore finding the best parameters for the Gaussian mixture.

Another statistical method that we almost finished implementing is multiclass SVM using *Error Correcting Codes (ECOC)*. The pro of this approach is that Matlab already implements direct support for multiclass SVM through ECOC, thus most of the work is already done. What we are trying to do is adapting our data and problem to Matlab's implementation and trying to optimize both the ECOC and SVM parameters.

3 Experiments and results

We will use the *Jim 2012 Chords* [jim] dataset, containing over 2.000 chords between minor and major triads with different musical instruments.

Our idea is to compare the results between the different approaches (the template-based and the statistical-based methods) as well as between different features (CENS, CLP, CRP), in order to have a 4 approaches by 3 feature matrix. Also, SVM could produce more results since so many parameters can be modified (kernel in primis). We will also probably use use different statistical measures such as precision, recall and F-measure.

References

- [chr] Chroma toolbox. <http://resources.mpi-inf.mpg.de/MIR/chromatoolbox/>.
- [Góm06] Emilia Gómez. Tonal description of polyphonic audio for music content processing. *INFORMS Journal on Computing*, 18(3):294–304, 2006.
- [jim] Jim 2012 chords. <http://www.montefiore.ulg.ac.be/services/acous/STSI/file/jim2012Chords.zip>.