COMBINING HOLISTIC AND ASPECT MODELS IN A CONTENT-DRIVEN MUSIC RECOMMENDATION SYSTEM

Stephan Hübler
TU Dresden, Germany
stephan.huebler@ias.et.tu-dresden.de

ABSTRACT

Music recommendation systems have become valuable tools for organizing music collections and discovering new music. This paper discusses our framework for a content-driven music recommendation system. We combine a holistic approach and classifiers trained to distinguish certain musical aspects to achieve optimal performance. The results of the different classifiers are combined via rank aggregation which allows for an easy extension of the framework by new modules. Holistic models and aspect classifiers use well established low level and mid level features.

This paper is part of the submission to the MIREX09 (Music Information Retrieval eXchange) audio music similarity and retrieval task.

1. OVERVIEW

In our framework, the process of generating music recommendations comprises three steps: (1) feature extraction from the audio signal, (2) extracting a model of the song, and (3) calculating similarities based on these models (see Figure 1). The audio file must be present in step (1).

The model generation step builds a compact representation of a song, a so-called *song model*, from various features extracted from the audio signal. This includes holistic features in a *bag of frames* representation as well as musical attributes using dedicated classifiers (cf. [1, 2]).

The third step generates a list of songs with attributes similar to a given query song. This list is aggregated from the result lists obtained by the classifiers using a music ontology. This aggregation technique ensures that the multi-dimensionality of music similarity is considered. The resulting recommendation list is sorted by decreasing similarity so that the most similar songs appear at the top of the list.

2. MODELING SONG CHARCTERISTICS

We apply two different approaches of modeling song characteristics in our system – holistic modeling and aspect classification. Both strategies generate result lists that are combined in a subsequent processing step using a priori information.

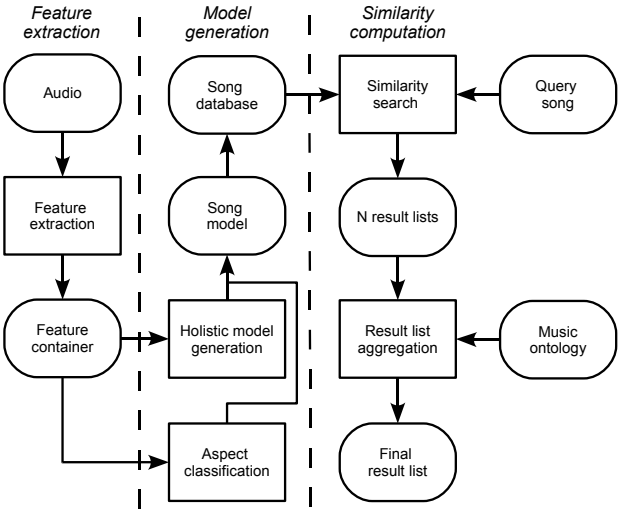

Figure 1. Flow chart of the recommendation process.

2.1 Holistic modeling

The holistic approach considers the whole song as a single entity. The distributions of selected low level features of a song are modeled and compared during retrieval, assuming that songs having similar feature distributions are regarded as similar. Table 1 lists features we use for holistic modeling. Merging the result lists generated by different holistic models moves those songs up the rank list that occur in most of the single result lists, hence yielding the consensus of all holistic models. Although holistic modeling tells us which songs are similar with respect to a given seed song, it doesn't tell us why or in which musical category the songs are similar.

2.2 Aspect classification

Every song has certain distinct characteristics like instrumentation, music color or sound density. To capture these characteristics, we train a Support Vector Machine (SVM) for every category listed in Table 2. Features are selected from the list of low level features and transformed into an appropriate feature space using Linear Discriminant Analysis (LDA). We noticed that the design of the traning sets

Feature	Dimension
Log Loudness	12
Norm Loudness	12
Mel Frequency Cepstrum Coefficients	16
Spectral Centroid	12
Spectral Crest Factor	16
Spectral Flatness Measure	16
Zero Crossing Rate	1
Onset Density	19
Excerpt of auto correlation function	70
Statistics derived from ACF	6
Percussiveness	19
ZCR Modulation	24
SFM Modulation	24

Table 1. Low- and mid-level features and their dimensionality.

is crucial for the performance of the aspect classifiers. Best results were obtained when the training data covered a broad variety in style of the selected musical characteristics.

During classification, the aspect classifiers generate a decision for every frame in the audio signal. Building a histogram from this result sequence provides an aspect profile of the song. This aspect profile represents the song in terms of the chosen musical attributes, e. g. whether the song is rather dark or bright in terms of music color (cf. Figure 2). Comparing the aspect profiles of the query songs with the profiles of all songs in the database using an appropriate distance measure (e. g. Common Area Norm) results in another set of similarity lists.

Aspect	Attributes
Mood	Happy, Sad
Music color	Dark, Bright
Sound density	Sparse, Full
Vocals	Singing, Instrumental
Instrumentation	Guitar, Piano, ...
Style	Rock, Electronics, Jazz, ...

Table 2. Musical aspect models.

2.3 Combining holistic and aspect classification results

During retrieval, both holistic modeling and comparison of aspect profiles from aspect classification yields result lists that need to be merged to a final result. We apply a set of rules represented in a ontology. Using an ontology is a convenient way of integrating information from different sources of information. Annotation information for the query song can be extracted from the aspect profile by selecting the attributes with the highest value in all the profiles. Provided that the automatic annotation is correct, we can assemble the final result list according to the characteristics of the query song. This offers more flexibility and optimization potential in combining holistic and aspect classification result lists. As an example, the weights for aggregating individual result lists might be different for classical query songs and rock query songs.

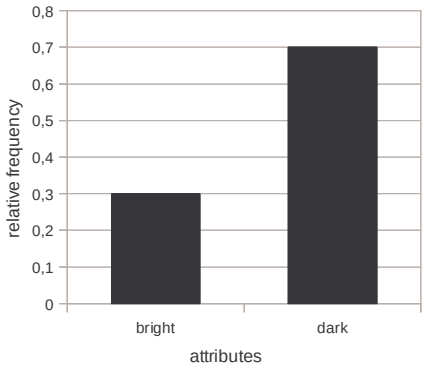

Figure 2. Example aspect profile for music color.

3. SUBMITTED SYSTEM

The submitted system supports three processing modes:

- Mode 1: holistic approach only
- Mode 2: holistic approach and aspect classification combined
- Mode 3: aspect classification only

4. FUTURE WORK

A prerequisite for the presented strategy of combining results from holistic and aspect classification is the reliability of the automatically generated annotations. We found in our experiments that this is an area with further optimization potential. Therefore, selecting additional musical aspects that cover the most important characteristics is one of our next steps. Another work in progress is to further develop and extend the ontology as well as to integrate user preferences and customization into the system.

5. REFERENCES

[1] Christoph Bastuck. An extensible and multiperspective approach for music similarity. In *3rd Music Information Retrieval Evaluation eXchange (MIREX 2007)*, Vienna, Austria, Sept 2007.

[2] Christoph Bastuck and Christian Dittmar. An integrative framework for content-based music similarity retrieval. In *DAGA 2008*, 2008.

[3] Stephan Hübler. Suchraumoptimierung zur Identifizierung ähnlicher Musikstücke. Master's thesis, Technische Universität Dresden, Fakultät Elektrotechnik und Informationstechnik, December 2008.