

Music Genre Classification

COS424

Vladimir Feinberg, Siddhartha Jayanti

9 February 2016

1 Abstract

blah blah [3].

2 Introduction

A basic problem in the field of music is to label a song into a *genre* (jazz, rock, country, classical etc.). In general, a given song can have multiple genres, and two people may disagree on whether a song s can be labeled with a given genre g . Furthermore, there are a huge number of genres and some are very similar to others. Thus, the most general version of the *genre classification problem* which intuitively asks for a complete list of genres given a song is hard to formalize and solve. In this paper, we will be discussing a simpler version of this problem which we call *Genre-Classify* that asks for a single genre label given a song. Our given data set contains 1000 different songs, which we shall call our song set S . Each song $s \in S$ is exactly 30 seconds long and is labeled with a single genre label $g(s) \in G = \{\text{blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, rock}\}$. There are exactly 100 songs of each genre. Our problem can be formally stated as follows: Given a random subset of pairs $P = \{(s, g(s)) | s \in T\}$ where $T \subsetneq S$, derive a function $g' : T' = (S - T) \rightarrow G$ that attempts to maximize the proportion of correct guesses $g'(s) = g(s), s \in T'$. In our experiments we let our *training set* T be picked as a random subset of size $0.8|S|$, and test the models g' we derive on our *test set* T' .

One problem in using algorithms on the music data, is the dimensionality of the data (if the song was interpreted as a vector of bits for instance). One way to reduce the dimensionality of the problem would be to take the first few components of a PCA; but better methods are possible when we have a labelled data set [4]. One such method is to encode musical *features* in a vector using the Fisher Linear Discriminant Analysis (Fisher-LDA) method [4]. This method is used on *timbre-related* musical features in Enrique et al. [1]. They obtain a 4.09% probability of error when they use the MFCC and Fisher-LDA. A later paper by Chang et al. [2] uses both short and long term features of music and the *compressive sampling* technique to keep the dimensionality low. They obtain 92.7% accuracy when they use multiple short and long term features.

References

- [1] ALEXANDRE, E., CUADRA-RODRIGUEZ, L., GIL-PITA, R., AND ROSA-ZURERA, M. Application of fisher linear discriminant analysis to speech/music classification. In *Audio Engineering Society Convention 120* (May 2006).
- [2] CHANG, K. K., SHING ROGER JANG, J., AND ILIOPOULOS, C. S. Iliopoulos: music genre classification via compressive sampling. In *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)* (2010), pp. 387–392.
- [3] DAVIS, M., AND LLEO, S. Fractional kelly strategies in continuous time: Recent developments.

[4] WELLING, M. Fisher linear discriminant analysis. unpublished note.