

Exercise 3 - Multimedia Data Mining

Group09

Samira Mahdi Zadeh- 1329639

Thiago Carneiro-1028762

Introduction

In this assignment, we performed a task in the Music Information Retrieval field, the audio music similarity and retrieval.

Dataset Preparation

For the dataset, we have downloaded and used the following collection:

http://kronos.ifs.tuwien.ac.at/ISMIRgenre_mp3_44khz_64kbit_mono_30sec.zip

This collection consists of 6 classes (genres):

- Classical
- Electronic
- Jazz_blues
- Metal_punk
- Rock_pop
- World

We used the Collection provided in the slide seeking for diversity in the collection, we used five very different genres:

Classical, Electronic, Jazz Blues, Metal-Punk and Rock-Pop.

Predominant characteristics of classical group is wide use of string instruments like violin, viola and cello and also Harpsichord and Piano as well as female voices, normally in highest pitch and men voices in low pitch.

In Jazz & Blues the use of woodwind/wind instruments such as saxophone and trumpet, and other instruments like drums, piano and guitar is predominant.

Electronic Music makes use of different electronic music instruments such as sound synthesizer and vocals (sometimes digitalized).

Metal & punk makes use of electric guitar, bass and drums as well as vocals (low in pitch and very loud).

Rock & Pop the prominent use of acoustic guitar, electric guitar, drums, synthesizer, keyboards and vocals is perceivable.

In World group many different traditional instruments which are associated to specific nations/countries (e.g. Indian, Persian instruments) are used.

There are some similarities between Metal and Rock groups, since the instruments used in both groups, such as electric guitar, are similar. But there are some differences which help to discriminate these two groups. One is the loudness; Metal songs are normally louder than Rock songs. Also the tempo of Metal song is more than Rock songs.

The approach to select the similarity and dissimilarity in each song group was based on the sound of the instruments.

Evaluation Preparation

Feature Extraction

For feature extraction, we use AudioFeatureExtraction software.

There were three features sets available: the Rhythm Histogram (RH), Rhythm Pattern (RP) and Statistical Spectrum Descriptor (SSD), which we focused on RH and SSD.

The collection that we use for our experiment consists of 30 second segments of songs in mono channel. Therefore no preprocessing was needed. These unified 30-sec segments will help find matches more precisely.

Perform similarity retrieval

We performed the similarity retrieval using SOMToolbox software

Evaluate and Summarize Results

The result is made by comparing two different feature sets (SSD and RH) and two different metrics. In summary we performed:

- 10 query songs with feature set RH, L2Metric;
- 10 query songs feature set SSD, L2Metric;
- 10 query songs feature set SSD, L1Metric

We evaluate the top 1-5th, 10th, 20th, 40th, 50th and 100th songs retrieved for the two first song

pairs of each genre, as shown in Table 1. For each query song, we assign a score from 0 to 2:

- 0 - Not similar
- 1 - Similar
- 2 - Very Similar

For having an overall view, we also sum up the grades for each evaluation and stored it in the last column of table.

Query	1	2	3	4	5	10	20	40	50	100	TotalScore
Classical 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	2 1 2	2 2 2	2 2 2	2 1 2	1 1 2	1 2 1	0 0 0	0 2 0	0 0 0	1 0 0	11 11 11
Classical 2 RH-L2Metric SSD-L2Metric SSD-L1Metric	1 2 2	2 2 2	1 2 2	2 2 2	2 1 1	0 1 2	2 0 0	0 0 1	0 1 1	1 0 1	11 11 14
Electronic 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	2 0 0	1 1 1	1 0 0	0.5 0 0	0 0 0	0 1 0	1 0 1	0 2 1	0 0 1	0 0 0	5.5 4 4
Electronic 2 RH-L2Metric SSD-L2Metric SSD-L1Metric	1 2 2	1 1 1	0 0 0	0 0 0	0 1 1	1 2 0	0 0 0	0 0 0	0 0 0	0 1 1	3 7 5
Jazz Blues 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	0 2 2	0 0 0	1 2 0	0 2 2	0 2 1	0 2 2	0 1 1	0 0 0	2 2 2	0 1 1	3 14 11
Jazz Blues 2 RH-L2Metric SSD-L2Metric SSD-L1Metric	2 2 2	1 2 1	0.5 0 2	1 1 0	0 1 0	0.5 0 0.5	0.2 0.5 0.2	0 0 0	1 _(same inst./style) 0.5 1	0 0 0.5	6.2 7 7.2
Metal Punk 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	2 2 2	2 2 2	2 2 2	1 1 2	1 1 2	1 1 1	1 1 0.5	1 0.5 0.2	1 1 1	0 _(classical?) 0 _(elec) 0.5 _(rock)	12 11.5 13.2
Metal Punk 2 RH-L2Metric SSD-L2Metric SSD-L1Metric	0 _(classic violin) 2 2	2 2 1	0 1 1	2 2 2	2 1 2	1 1 1	1 1 0	0 1.5 1.5	0 0.7 1	0 0 1	8 10.2 12.5
Rock Pop 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	2 0.4 0.4	2 0.4 0.4	2 0.4 0.4	1 0.4 0.4	1 0.4 0.4	0.4 0.4 0.4	1 0.4 0.4	1 0.4 0.4	0.4 0.4 0.4	0.2 0.4 0.4	11 4 4

Rock Pop 2											
RH-L2Metric	1	2	1	1	0	0	0.5	0.5	0	0	6
SSD-L2Metric	2	1	1	0.5	0.5	1	0.5	0.5	0	0	7
SSD-L1Metric	2	2	2	1.5	1.5	0.5	0.5	0	0	0	10

Table 1: Evaluating the top 1-5, the 10th, 20th,40th,50th,100th song.

Our observation

From “Total score” column in Table 1 it is perceivable that, comparing to other groups, both RH and SSD perform relatively very well on Classical genre, and retrieved a lot of Classical songs, very near to the query song..

After that, the Metal-Punk group is the next genre that RH and SSD have a good performance on it.

But neither SSD nor RH performs well for Electronic genre, retrieving mostly songs, very far from the query song.

For Jazz-Blues, SSD performs relatively better than RH. RH had a poor performance for the first Jazz Blues song, retrieving a bunch of Metal and World songs.

The song representing Rock-Pop 1 is an exceptional case, which were interesting to evaluate. Although it is categorized in Rock and Pop group but it has special characteristics which makes it different: there is no beats introduced by drums, not many musical instruments, there is only a continuous sound without pause from beginning to the end. The Similar Songs retrieved by RH were mostly from classical genre and most of them were with only a choir of men and no instruments. So we could consider them as similar match to the original song. But the SSD on the other hand, had a different approach for detecting similar song. The similar song from SSD perspective were purely a single instrument (=harp) playing a melody. From our point of view RH results were more similar to the original song than SSD's.

Considering L1Metric and L2Metric, their performance is close to each other with maximum difference of 3 in their total score.

Table 2 shows the location, at which the second song of the similar/dissimilar pairs is.

query	Similar Song Location in Result List	Dissimilar Song Location in Result List
Classical 1		
RH-L2Metric	83	
SSD-L2Metric	29	
SSD-L1Metric	12	

Classical 2 RH-L2Metric SSD-L2Metric SSD-L1Metric		Not in first 100 result (>100) Not in first 100 result (>100) Not in first 100 result (>100)
Electronic 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	Not in first 100 result (>100) 28 40	
Electronic 2 RH-L2Metric SSD-L2Metric SSD-L1Metric		Not in first 100 result (>100) Not in first 100 result (>100) Not in first 100 result (>100)
Jazz Blues 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	Not in first 100 result (>100) 1 1	
Jazz Blues 2 RH-L2Metric SSD-L2Metric SSD-L1Metric		Not in first 100 result (>100) Not in first 100 result (>100) Not in first 100 result (>100)
Metal Punk 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	9 3 3	
Metal Punk 2 RH-L2Metric SSD-L2Metric SSD-L1Metric		Not in first 100 result (>100) Not in first 100 result (>100) Not in first 100 result (>100)
Rock Pop 1 RH-L2Metric SSD-L2Metric SSD-L1Metric	21 Not in first 100 result (>100) 39	
Rock Pop 2 RH-L2Metric SSD-L2Metric SSD-L1Metric		Not in first 100 result (>100) Not in first 100 result (>100) Not in first 100 result (>100)

Table 2: Location of similar/dissimilar songs in Result List.

From Table 2 it is perceivable that:

- Since none of dissimilar songs are within top 100 results returned by the software, we as humans are in strong agreements with the machine perception about dissimilarity of songs.
- In general (except one case) the Location of similar song using SSD feature set is more close to top of ranking list than the RH.

Query Name	Song
Classical 1 (1=taken from similar set)	classical_artist_15_album_2_track_4
Classical 2 (2= taken from dissimilar set)	classical_artist_10_album_1_track_2
Electronic 1	electronic_1-dezert
Electronic 2	electronic_8-ehaw
Jazz Blues 1	jazz_blues_18-slide_boogie
Jazz Blues 2	jazz_blues_3-my_momma_told_me
Metal Punk 1	metal_punk_6-in_my_livid_eyes
Metal Punk 2	metal_punk_7-heavenly_rain
Rock Pop 1	rock_pop_3-air_chair
Rock Pop 2	rock_pop_12-sympathetic_stranger

Table 3: Mapping between Query names used in this report and the song titles.