Guitar Type Classification

https://github.com/sergionimar/MIR-Project_SergioNieves

Sergio Nieves Martínez

This project involves the development of a software that can distinguish between recordings of songs that are played with an acoustic guitar or with an electric guitar without any kind of distortion. To do this I have used the IDMT-SMT-GUITAR database, which can be found at the following link: https://www.idmt.fraunhofer.de/en/business_units/m2d/smt/guitar.html. Concretely I have used the fourth dataset of which the database is composed, which contains a great number of songs recorded by 3 guitars (two electric and one acoustic played in two different ways) and which is divided inside each guitar by the speed of the song and by the musical genre.

To create the software, I have based myself on two papers, which have helped me choose the machine learning model to use and the features necessary to classify these two types of guitars. The papers are the following ones:

- Johnson, D., & Tzanetakis, G. (2015). Guitar model recognition from single instrument audio recordings. *IEEE Pacific RIM Conference on Communications, Computers, and Signal Processing - Proceedings*, *2015–Novem*, 370–375.
- Fuhrmann, F. (2012). Automatic musical instrument recognition from polyphonic music audio signals.

For the creation and evaluation of the Support Vector Machine algorithm I have used the sklearn python library, and for obtaining the sound features of the recordings, the Essentia library developed by the MTG.

The objective of this project is basically to find the right way to train the machine learning model so that it works with the greatest possible accuracy. To do this I have been exploring several ways to group the files available for both training and testing.

First, I tried to train the algorithm with the slow tracks and test it with the fast ones. The features of the songs that I have used in this case were MFCC only, computed for 1 second frames with an overlapping of 50%. The results by proving both with genres individually and together (*Table 1*) were about a 55% of accuracy. This is a very bad data considering that we only have two possible results, acoustic or electric. Also, in all the cases practically all the guitars were detected as of only one type, all acoustics or all electrics.

```
Accuracy of the model:  0.539421709591

Confusion matrix:
Acoustic Electric
  3022      196   Acoustic
  2719      392   Electric
```

Table 1: results using slow tracks for training and fast tracks for testing. Features: MFCC only

The next step was to try different sets of features. We extracted the features that are supposed to be most relevant to distinguish between the two types of guitar, MFCC, energy in Bark Bands and the 0th spectral valley coefficient. We discovered that the best option was to use them all together, but the result of training and testing with fast and slow songs was still very bad, so we decided to use other combinations of tracks. We stopped differentiating between fast and slow songs and started using all the songs of the same genre together. We use sets of different genres for training and testing and tried to balance the genres that made all guitars be detected as electric and those that made all be detected as acoustic. The best accuracy was obtained using the genres classic, jazz, pop and metal for training and reggae_ska, country_folk, rock_blues and latin for testing (*Table 2*). We obtain an accuracy of 69% that improves considerably the previous result but that continues to wrongly classify more than half of the acoustic guitars, while the electric guitars are correctly classified.

```
Accuracy of the model:  0.693103951013

Confusion matrix:
Acoustic Electric
  1563    2046   Acoustic
  59      3191   Electric
```

Table 2: results using classic, jazz, pop and metal for training and reggae_ska, country_folk, rock_blues and latin for testing. Features: MFCC, energy in Bark Bands and the 0th spectral valley coefficient

Finally, following the recommendations of my classmates during the presentation of the project I decided to change the configuration of the algorithm created by the svc function of the sklearn library. Just by changing the kernel type from 'rbf' to 'poly' we can improve the effectiveness of the model up to 87% accuracy (*Table 3*).

```
Accuracy of the model:  0.869806094183

Confusion matrix:
Acoustic Electric
  3541     68   Acoustic
  825      2425   Electric
```

Table 3: Results using classic, jazz, pop and metal for training and reggae_ska, country_folk, rock_blues aand latin for testing. Features: MFCC, energy in Bark Bands and the 0th spectral valley coefficient. Kernel: 'poly'

With this model we already get quite good results, but there is still a small problem, in this case the algorithm will work much better detecting acoustic guitars than electric ones. Adding more spectral valley coefficients (the first five) to the algorithm, we get 85% accuracy but with a more balanced performance between both types of guitar (*Table 4*).

```
Accuracy of the model:  0.852165038635

Confusion matrix:
Acoustic Electric
  3133     476   Acoustic
  538      2712   Electric
```

Table 4: Results using classic, jazz, pop and metal for training and reggae_ska, country_folk, rock_blues aand latin for testing. Features: MFCC, energy in Bark Bands and the 0 to 5 spectral valley coefficient. Kernel: 'poly'

We have to keep in mind that these results are for each analysed frame of each song used for testing. The final result to decide the instrument found in a song will be made by majority voting between the frames that compose the song, so it is expected that the result improves if most of the frames of a song are well classified. For this reason, we consider that the best option is to use this last algorithm, since the probabilities that both guitars after majority voting were well classified is greater.