Symbolic Music Similarity and Music Plagiarism

Paul Kröger Ali Bektas

November 30, 2019

Abstract

As it becomes ever more easy to produce music from one's own home, the challenge of detecting plagiarized parts becomes more important. Detection of such plagiarism through subjective reasoning is a time consuming effort and must inevitably be replaced by automatized solutions.The problem of detecting such plagiarisms is strongly connected to determining similarity between two pieces by means of music similarity approaches. Thus , the detection of plagiarized works is based on techniques of music similarity. Not every approach in music similarity is however useful , since music similarity is a broader term than fetching related songs based on their melody , rhythm and so on. Due to melodic similarity being the crucial factor in plagiarism cases [2] , and rightly so since the melody is according to [**?**] the most important aspect in a song for the human ear, we will focus on melodic similarity detection . The main emphasis of this paper is therefore to explore the ways in deciding whether two pieces are similar, what metrics can be used to do so and how appropriate the different approaches are.

Outline

Introduction

Plagiarism detection algorithms differ from music similarity algorithms in that they answer a Yes/No question whereas music similarity algorithms return a list of best results.

1 Approaches for symbolic data

Symbolic data in music are representations of a musical structure usually based on pitch related and temporal informations. Modern western techniques date back to ...

In modern digitalized techniques used in representing a musical piece there are more than the two abovementioned dimensions.

Rhytmical structure is as important since depending on the rhythm , notes become more significant within a piece.

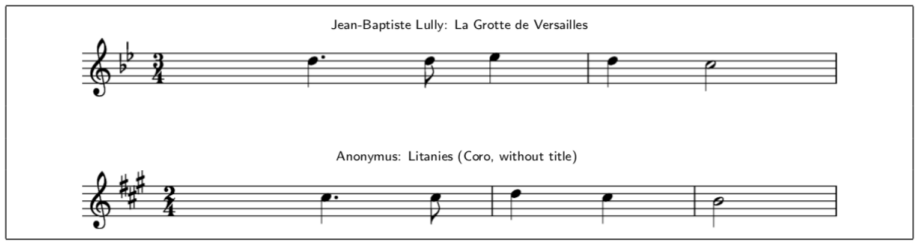

Figure 1: Even though the notes are the same they are heard different since rhythm changes how a melodic sequence is percieved.

1.1 Symbolic Music Similarity

The MIREX (Music Information Retrieval Exchange) is a community who arranges contest in different topics of Music Information Retrieval. Since 2005 there has been a number of algorithms which have suggested new approaches in the field. [3] places these algorithms with into three categories: The ones that use

1. Mathematics
2. Cognition
3. Music Theory

1.1.1 Algorithms of Category 1

The algorithm proposed by [1] aims to convert melodies into "orthagonal polygonal chains" and then "find the minimum area of minimum area between two lines and two horizontal edges".

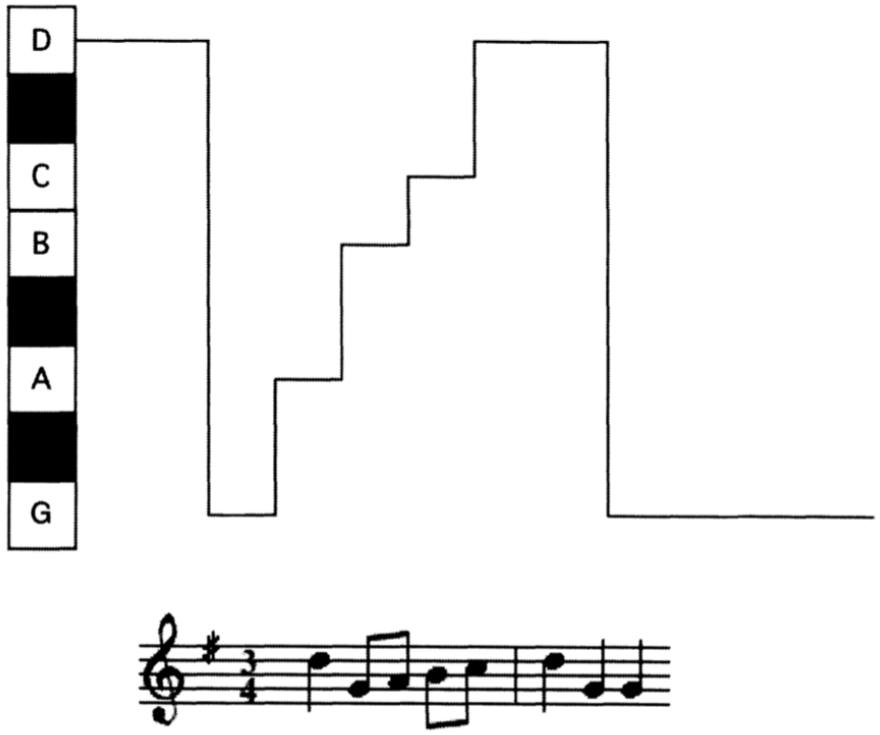

Figure 2: Source: [1]

1.1.2 Algorithms of Category 2

An algorithm proposed by [5] uses the Implication/Realization Model (IMR). We will discuss IMR.

A MIDI-formatted input undergoes Melody Seperator , which as the name suggests turns melody into segments. This paper emphasizes that the intervalic information between two notes must be further detailed by the direction of pitch change.

This algorithms incorporates n-gram method. What is n-gram method?

Similarity is finally calculated by

$$S(query, comparison) = \frac{n}{max(a, b)} \tag{1}$$

Where a and b are reprectively element count of query melody and melody compared.

1.1.3 Hybrid Algorithms

Hybrid algorithms are the algorithms , that it incorporate all of the above-mentioned (See 1.1) categories. This is done by taking a linear combination of several single algorithms.

The SIMILE [4]: The algorithms that are used

- Edit Distance: e.g. Mongeau & Sankoff, (1990)
- n-grams: e.g. Downie (1999)
- Geometric distance: Steinbeck (1982);O'Maidin (1998)
- Correlation coefficient: e.g. Steinbeck, (1982)

$$\begin{aligned} optP = 0.31 \cdot joint412 + 1.37 \cdot rawEd \cdot connEd(1) + 0.643 \cdot rawEd \cdot harmCoEA \\ \text{–}1.55 \cdot connEd \cdot bGrCoorF \\ + 0.65 \cdot bGrCoorF \cdot rhytFuzz \\ \text{–}0.39 \cdot harmCoEA \cdot rhytFuzz\text{–}0.133 \end{aligned}$$

We will explain what these are.

2 Approaches for plagiarism detection

Here we will introduce different approaches to automatically detecting plagiarisms in music.

2.1 Fuzzy vectorial-based similarity detection

An algorithm proposed by [6] that constructs a vectorial representation of a melody consisting of its intervalic changes and the corresponding rhythm. These representations are then used to calculate a fuzzy number to determine similarity between melodies. Vectorial melodic distance between two subvectors of length k is then calculated by the following formula :

$$D(s1, s2) = (D_m(s1, s2), D_r(s1, s2)) \tag{2}$$

where

$$D_m(s1, s2) = \sqrt{\sum_{i=1}^{k} |n_i * V_{M_1} * getPos(n_i) - m_i * V_{M_2} * getPos(m_i)|^2} \tag{3}$$

and

$$D_r(s1, s2) = \sqrt{\sum_{i=1}^{k} |d_i * R_{M_1} * getPos(d_i) - t_i * R_{M_2} * getPos(t_i)|^2} \tag{4}$$

The vectorial melodic distance is the used to calculate a Fuzzy Number. We will explain how this process works.

Conclusion

In this paper we have reviewed various algorithms for measuring symbolic melodic similarity and algorithms for detecting music plagiarism...

A pair of sentences about why plagiarism being a Yes/No question makes trouble since general algorithms in music similarity are not yet well established to be able to provide a unified solution to music of all genres.

References

[1] G.Aloupis,T.Fevens,S.Langerman,T.Matsui,A.Mesa, Y. Nufiez, D. Rappaport, G. Toussaint. Computer Music Journal , 30:3, pp. 67-76, Fall 2006.

[2] Cronin, C. 1998. "Concepts of Melodic Similarity in Music-Copyright Infringement Suits."InW.B.Hewlett and E. Selfridge-Field,eds. Melodic Similarity: Concepts, Procedures,and Applications. Cambridge, Massachusetts:MITPress,pp. 187-209.

[3] Symbolic Melodic Similarity: State of the Art and Future Challenges Valerio Velardo,Mauro Vallati and Steven Jan Computer Music Journal, 40:2, pp. 70–83, Summer 2016 doi:10.1162/COMJ a 00359 2016 Massachusetts Institute of Technology.

[4] Velardo, Valerio, Mauro Vallati, Steven Jan. "Symbolic melodic similarity: State of the art and future challenges." Computer Music Journal 40.2 (2016): 70-83.

[5] Sakurako Yazawa, Yuhei Hasegawa, Kohei Kanamori and Masatoshi Hamanaka Melodic Similarity based on Extension Implication-Realization Model Music Information Retrieval Exchange , 2013.

[6] R. D. Prisco, D. Malandrino, G. Zaccagnino, and R. Zaccagnino, "Fuzzy vectorial-based similarity detection of music plagiarism," in 2017 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2017, Naples, Italy, July 9-12, 2017, 2017, pp. 1–6.