A Study on the Measurement of Music Works and Genres Similarity

Wei Yang [1, *], Linsheng Hu [1], Xinyu Si [1]

[1]School of Finance, Xi'an Eurasia University, Xi'an 710065, China

*Corresponding author

Keywords: music works music genre similarity euclidean distance contour coefficient pca

Abstract: In this paper, we measure the similarity between musicians and other genres, and show that the influence of different genres is the distance between musicians and other genres. Finally, the similarity between different genres is analyzed by PCA dimensionality reduction visualization, which shows the dynamic change of similarity between different genres.

1. Introduction

Music is one of the most important art forms of human civilization so far. Music is also a part of human society and an important manifestation of culture. In the early world civilization, since the community established a relatively complete social system, it set up special departments and individuals in charge of music. After evolution, music has been divided into different schools and created more musical features. Every musician in the process of creation will be influenced by the predecessors, many songs have similar characteristics (intensity, amplitude, beat, etc.). Will musicians be influenced by their predecessors in the process of creation? Will their works have a certain degree of similarity? If so, are the works created by musicians of the same genre the most similar? In view of the above problems, this paper will use the method of machine learning to peel the cocoon layer by layer between music works, and build different models to measure the similarity between music works and genres.

2. Data Collection and Processing

In order to study the similarity of different artists' music works, the music works of different musicians from 1930 to 1980 were collected. After processing, each music work was represented by the indicators of duration, energy, beat and rhythm. In order to compare the similarity of each musician's works, the average value of each musician work on each index is calculated. Then all the indicators are standardized for subsequent calculation.

3. Similarity Measurement of Musicians' Works Based on European Distance

3.1 Euclidean Distance Principle

European distance is used to measure the similarity between musicians' works. The formula of Euclidean distance is as follows:

$$dist(A, B) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2}$$

Where a and B represent two musicians, a and B respectively_ i,B_ I is the corresponding index. The smaller dist (a, b), the more similar the music of the two musicians.

3.2 Similarity Measurement Results of Musicians' Works

By calculating the Euclidean distance, the top three groups of musicians with the shortest Euclidean distance are found, as shown in Table 1

Table 1 Most Similar Three Pairs of Musicians

Musician 1 name	Musician 2 name
Usher	Sammie
The Beatles	Matt Simons
D'Angelo	J. Holiday

At the same time, compare the music characteristics of the three groups of musicians, as shown in Figure 1

According to figure 1, we can see that the duration, speech and energy of the works of similar musicians are similar. Then we compare the names of artists in Table 1 with the influence form of musicians who have mutual influence, and find that there is mutual influence between the most similar musicians.

4. Similarity Measurement between Musicians and Genres Based on Silouette Coefficient

4.1 Silhouette Coefficient Principle

Based on the calculation of Euclidean distance, this paper uses the contour coefficient to measure whether musicians of the same genre are more similar than musicians of different genres.

The formula for calculating the contour coefficient of each musician's work is as follows:

$$s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \qquad s(i)= \begin{cases} 1 - \frac{a(i)}{b(i)}, & a(i) < b(i) \\[2mm] 0, & a(i) = b(i) \\[2mm] \frac{b(i)}{a(i)} - 1, & a(i) > b(i) \end{cases}$$

Here a (I) is the average distance between the i-th musician's work and other musicians' works of the same genre. The smaller the a (I), the more the i-th musician belongs to this genre. B (I) is the minimum of the average distance from the i-th musician's work to all other genres. The larger the B (I), the less the i-th musician belongs to other genres.

Here a (I) is the average distance between the i-th musician's work and other musicians' works of the same genre. The smaller the a (I), the more the i-th musician belongs to this genre. B (I) is the minimum of the average distance from the i-th musician's work to all other genres. The larger the B (I), the less the i-th musician belongs to other genres.

4.2 Similarity Measurement Results of Works and Their Schools

Figure 2 is obtained by calculating the contour coefficient of each musician's work. As can be seen from Figure 2, the value of outline coefficient of many musicians' works is close to 0. It shows that many musicians are at the boundary of genres and may be influenced by many genres. At the same time, the contour coefficient of some musicians' works is greater than 0.6, which indicates that this musician is more similar to other genres. Musicians more similar to other genres are shown in Table 2.
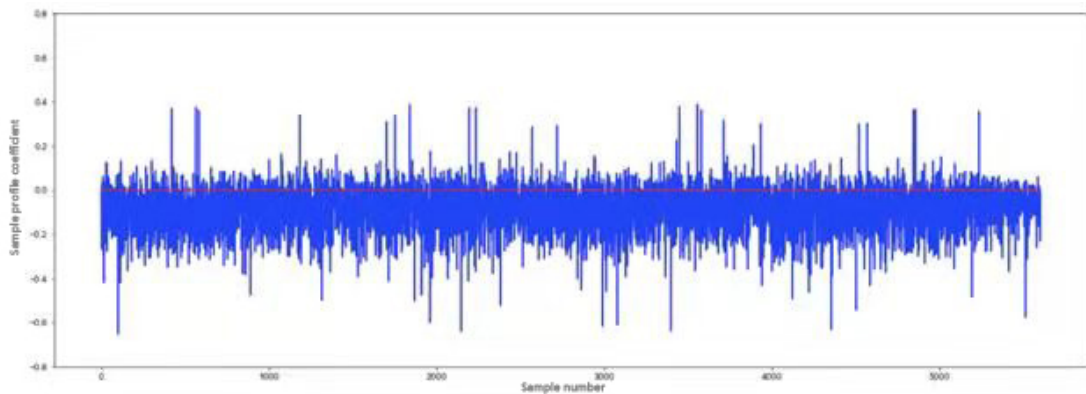

Fig.2 Figure of Contour Coefficient of Each Artist's Work

Table 2 Ranking of Musicians More Similar to Other Genres

artist_name	Silhouette Coefficient
Spinal Tap	-0.6533218
Jerry Lewis	-0.6415549
Tiny Tim	-0.639292
Cassius Clay	-0.6326768
Monty Python	-0.6181497
Flight of the Conchords	-0.6109727
Robert Klein	-0.6000388

From the results of Figure 2 and table 2, some musicians are also influenced by other schools of

music. Compared with the schools belonging to them, their music works are more similar to other schools.

5. Comparison of Music Genre Similarity Based on Pca

In order to further study the similarity between music genres, based on the known main genres of each musician, the dimension reduction method is used to visualize the relationship between different genres.

5.1 Comparison of Similarities between and within Schools

In this paper, PCA dimension reduction method is used to project multiple indicators of the works onto the two-dimensional plane, and color is used to distinguish different genres to visually show the similarity between and within genres,

Fig.3 Cluster Diagram of Similarity between Schools and within Schools
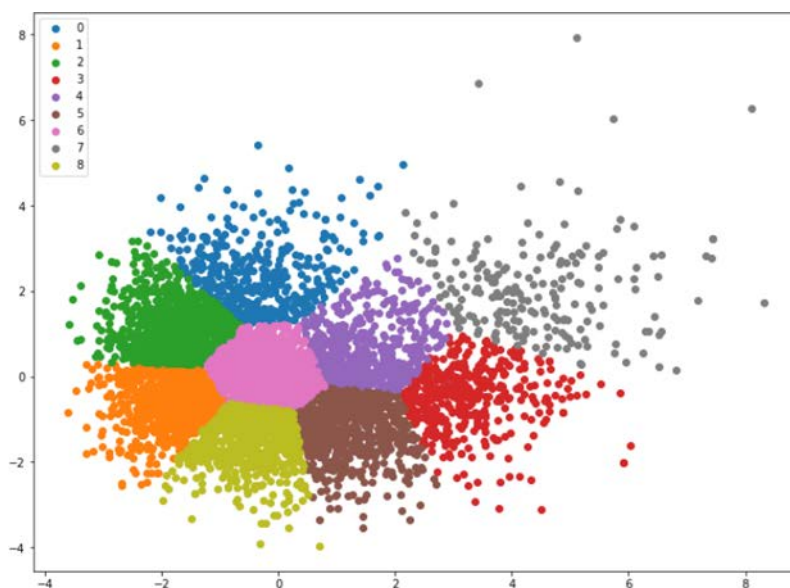
In the figure, 0, 1, 2, 3, 4, 5, 6, 7 and 8 represent easy listing, R & B, reggae, vocal, fold, classic, pop / rock, country and blues respectively. It can be seen from Figure 3 that pop / rock schools are relatively concentrated, which indicates that this school has a high degree of similarity, while reggae and country schools are relatively scattered and adjacent, which indicates that there is a great similarity between the two schools.

5.2 Changes of Music Genre Similarity over Time

In order to further analyze the changes of genre similarity with time, the data is divided into sections according to the year, with every 10 years as a span, and then the PCA dimension reduction visualization is carried out respectively, and Figure 4 is obtained. It can be seen from the results in Figure 4 that there are differences in the similarity between music genres in each stage.
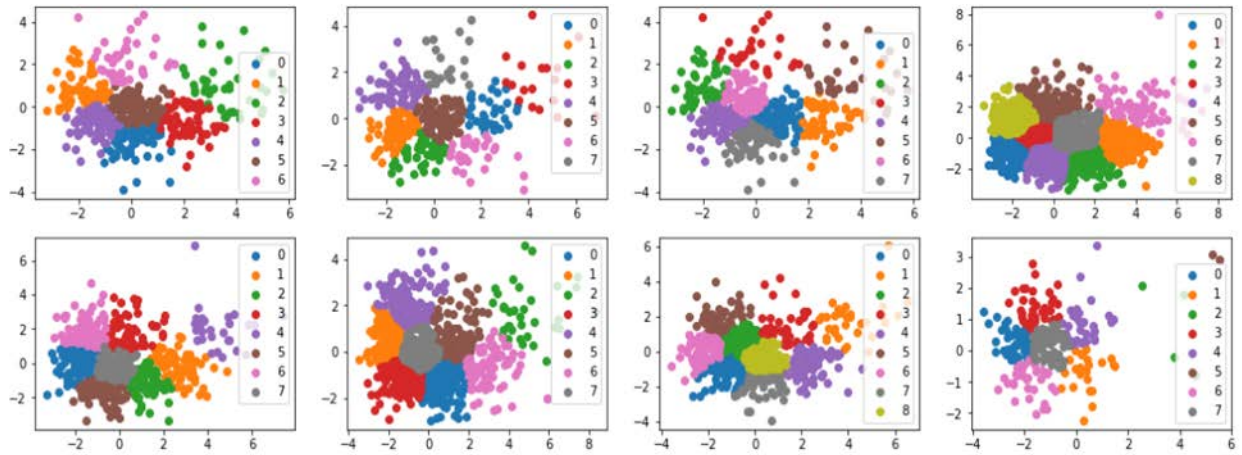
Fig.4 Cluster Diagram of School Similarity and Influence (1930-2010)

Figure 4 shows that the characteristics of music schools are changing dynamically, and the degree of similarity between music schools in different periods is also changing dynamically. Music itself is constantly evolving and updating with time and social environment.

6. Conclusion

This paper uses machine learning method to measure the similarity of different musicians' works, and measures whether musicians are more similar to other genres. It shows that there is influence between musicians, and there is cross genre influence. At the same time, it analyzes the similarity between different genres, which shows the dynamic change of similarity between different genres.

Acknowledgement

References

[1] Founder of Hungarian Conservatory of music, Tianjin, 2011

[2] Li Yang, Wu Yadong, Liu Baolong. A new approximate melody matching method and its application in humming retrieval system [J]. Computer research and development, 2003 (11): 1554-1560

[3] Ge Liang, Lang Jiangtao, Tang Huang, Tang Yunheng. PCA hubness clustering method for high dimensional data [J]. Modern computer (Professional Edition), 2017 (11): 52-55 + 59

[4] Lang Jiangtao. Research on PCA hub clustering method for high dimensional data [D]. Chongqing University, 2017