ISSN 2618-1568 Vol. 3, Issue 2: 60-64, DOI: 10.25236/FAR.2021.030212

Research on Music Influence Evaluation System Based on Similarity Measure and Grey Relational Analysis

Zhen Xiong, Chao Li, Junkang Jiang, Caiying Zhou, Qiang Wang

College of Science, Jiangxi University of Science and Technology, Ganzhou, Jiangxi, 341000, China

Abstract: *Music is an important part of the spiritual home of human beings, and it influences every generation with its own unique advantages. This paper explored the influence of music through the establishment of a directed network of music influence and a similarity measurement model, studied the evolution of music, and analyzed the influence process of the evolution. The research in this article explained the influence of music, the influence relationship between musicians and the evolution of music, which could make music develop better. We needed to solve the problems of artist similarity, genre similarity, and musical feature appeal. In this regard, we built a similarity evaluation model, read the music data set, and selected data for the indicators. We used principal component analysis to reduce the dimensionality of the index data, and then calculated the similarity based on Euclidean distance. For artist similarity, we divided it into the same genre and different genres, and we calculated that artists of the same genre were more similar than artists of different genres. For genre similarity, we calculated and drew a heat map of genre similarity. For the appeal of music features, we calculated the correlation between each feature and the popularity index based on the gray correlation analysis, and concluded that valence is the most infectious. This paper made full use of the music data obtained, described and quantified the influence of music from multiple aspects. The model in this paper is reasonable and realistic. The model can also be used to measure the impact of other occupations in society, and has a certain generalization.*

Keywords: *Similarity Measurement, Grey Relational Analysis, Influence Analysis*

1. Introduction

Music has been a part of human society since the beginning of time as an impor-tant part of human spiritual homeland [1]. It has its own unique advantages affecting generations of people. When artists create a new piece of music, there are many fac-tors that affect them, such as their innate originality, current social or political events, opportunities to acquire new instruments, or other personal experiences. Past music will also have an impact on new music and music artists, which is often reflected in the similarity of song structure, rhythm or lyrics. Many artists have contributed to major changes in music genres.

Many songs have similar melody. By considering song network structure and musical characteristics, music could be developed better.

2. Similarity measurement model

The similarity measurement model mainly measures the issues of artist similarity, genre similarity and genre appeal. Its general idea diagram is as follows:

ISSN 2618-1568 Vol. 3, Issue 2: 60-64, DOI: 10.25236/FAR.2021.030212

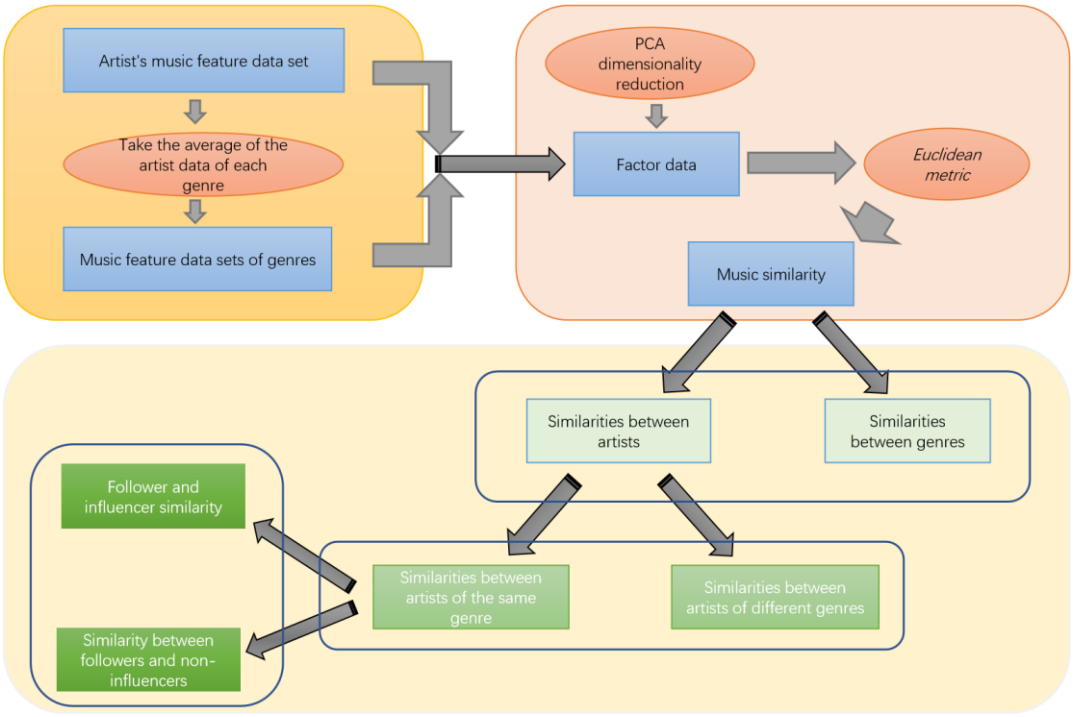


Figure 1: Conception of Similarity Measurement Model

Artist's similarity

According to the genres of the artists, the artists in the data_by_artist.csv file are divided to filter out artists of the same genre and artists of different genres. Since the artist_id feature has little effect on artist similarity, we removed the artist_id attribute from the filtered data. In order to eliminate the dimension, we normalized the data.

$$X_i' = \frac{X_i}{\sum_{i=1}^{n} X_i}$$

Where n is the total number of artists in the same genre, X_i is the feature value of the audio, and X_i' is the normalized feature value. After that, we used principal compo-nent analysis(PCA) [2] to reduce the dimensionality of the audio features, and replace the original more variable indicators with several comprehensive indicators, and reflect as much information as possible.

Finally, we reduced 14 features to 6 features through principal component analysis, which can save more than 90% of the information, indicating that the dimensionality reduction effect is better.

Euclidean distance is a commonly used distance definition, which refers to the true distance between two points in m dimensional space. The similarity can be regarded as the true distance between two artists in the 6-dimensional audio feature space, so the Euclidean distance can calculate the similarity of the artists well. The formula for calculating Euclidean distance is as follows:

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_m - y_m)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Where m represents the dimension of the audio feature, n=6.

We randomly select data from artists of the same genre and artists of different genres, and use Euclidean distance to calculate similarity based on normalization and principal component analysis. Define the distance of artists of the same genre as SameArt, and the distance of artists of different genres as DiffArt.

$$\text{SameArt} = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}(x,y), \text{Diff } Art = \sum_{i=1}^{n} \sum_{j=1}^{n} d'_{ij}(x,y)$$

Here, we select 17 artists of the same genre and different genres, that is, n=17.By calculating SameArt

ISSN 2618-1568 Vol. 3, Issue 2: 60-64, DOI: 10.25236/FAR.2021.030212

as 1272.40 and DiffArt as 1285.96, so the distance between artists of the same genre is smaller than the distance between artists of different genres. It shows that the differences between artists of the same genre are smaller than those of different genres. In other words, artists within genres are more similar than artists between genres.

Genre's similarity

In order to measure genre similarity, we need to obtain genre data. Each artist has a genre, we connected the artist to the corresponding genre in the data_by_artist.csv data set. Then, we grouped the data according to the differences in genres, used the mean to calculate, and used the results as the characteristics of each genre. Next, we drew the following radar chart, as shown in Figure 2, which visually illustrates the differences in different characteristics of each genre.

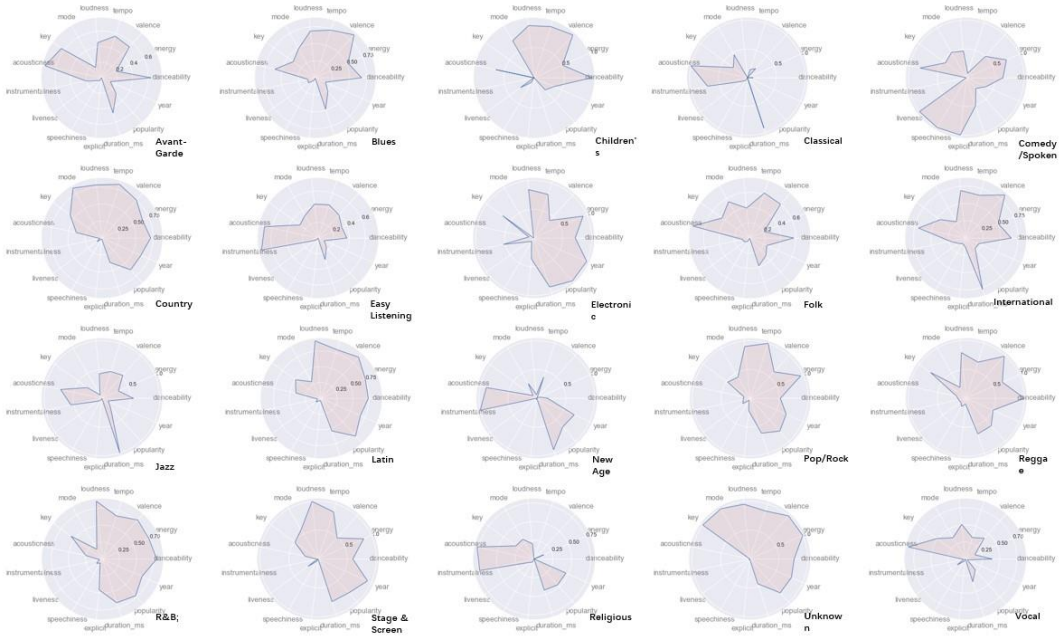


Figure 2: Radar plot of detailed performances indicators of music for 20 genres

Impact verification

In order to eliminate the influence of the dimension, we performed normalization, using Euclidean distance to measure the music similarity between followers and their influencers and the music similarity of non-influencers. For example, we selected the Pop/Rock genre, obtained a certain influencer (such as *NSYNC) from the influencer_name indicator, and found out the affected followers and the unaffected followers. We calculated the Euclidean distance between followers and their influencers as 1697.72, and the Euclidean distance be-tween followers and influencers as 1712.49. This shows that in the same genre, the songs created by followers are more similar to influencers than non-influencers, and the works created by followers will be affected by influencers.

Through the previous analysis and calculations, we have verified that followers will be affected by influencers. Next, we did further analyze the music characteristics to find out the "infectious" characteristics. The "infectiousness" of music is often inseparable from the degree of popularity, and more popular music tends to be more "infectious." In order to explore which music characteristics have a greater impact, we used grey relational analysisto measure the relationship between music characteristics and popularity. Grey relational analysis [3] refers to the quantitative description and comparison method of the development and change of a system.If the trends of the two quoted changes are consistent, that is, the degree of synchronization is higher, it can be said that the two have a higher degree of correlation.

3. Model solving

Euclidean distance can reflect similarity very well, so we also used Euclidean distance to measure the similarity between genres. We brought in the characteristic data of the genre, calculated the "distance" of different genres, and drew a heat map(shown in Figure 3).In Figure 3, the value represents the "distance"

ISSN 2618-1568 Vol. 3, Issue 2: 60-64, DOI: 10.25236/FAR.2021.030212

between the genres, the smaller the value, the higher the similarity of the genre, and 0 represents the same genre.

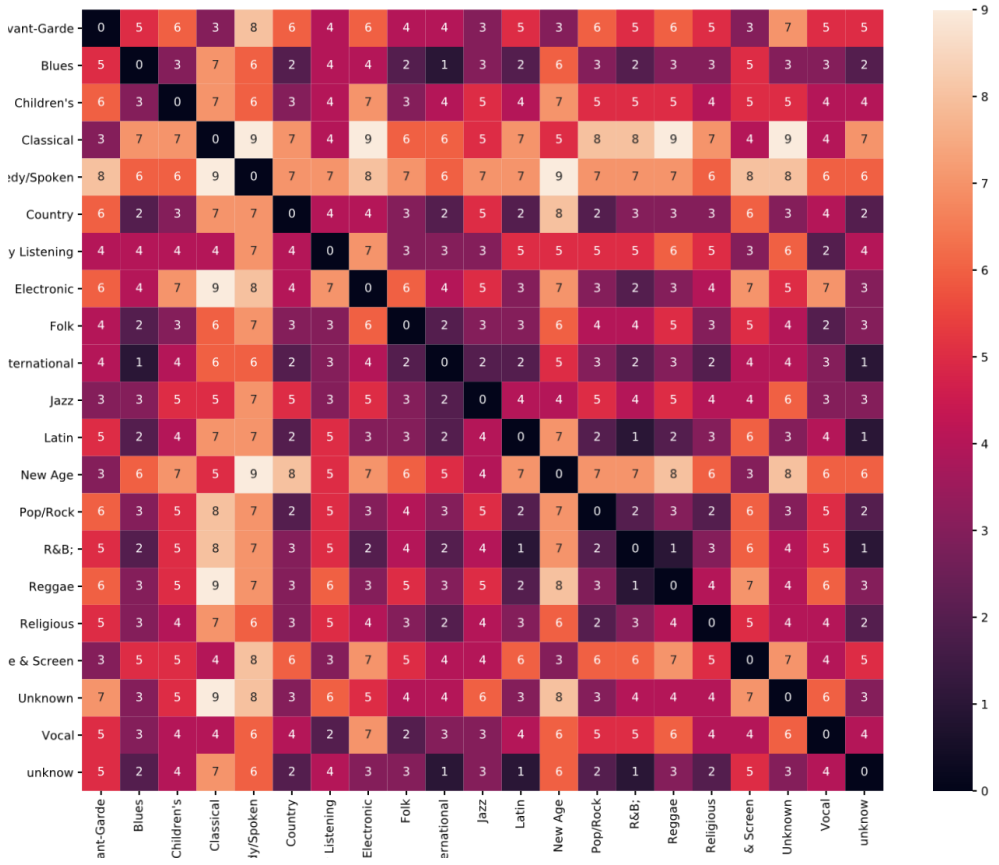
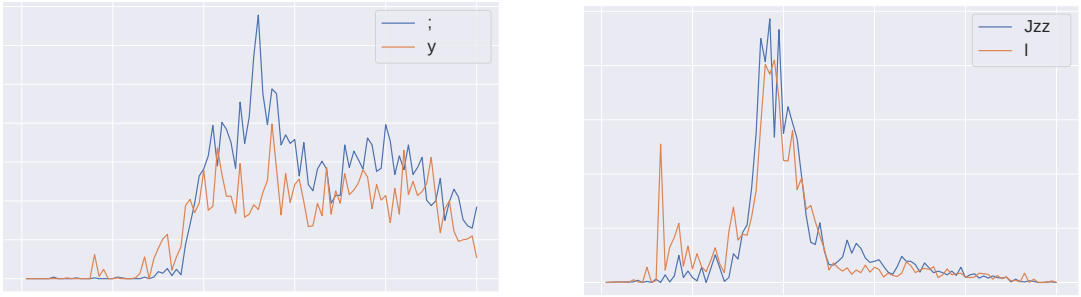


Figure 3: Genre indicator heat map

It can be seen from the above figure that the Blues genre is highly similar to the International genre; Latin and R&B are highly similar; R&B and Reggae are highly similar, with a value of 1.

In order to research the development of genres, we extract the data of each genre in full_music_data.csv, and then count the number of releases of each genre over time. Then, we conducted time series analysis on the statistical data, and visualized it, and drew a trend graph of the number of releases and popularity of each genre with the year.



(a) R&B and Country (b) Jazz and Vocal

Figure 4: Patial genre development trends

4. Conclusion

We adopted the similarity evaluation model, reduced the dimension by principal component analysis, and used Euclidean distance as the similarity measure to obtain that artists of the same school and different schools are more similar. Grey correlation analysis is used to find the most infectious characteristic of music.

ISSN 2618-1568 Vol. 3, Issue 2: 60-64, DOI: 10.25236/FAR.2021.030212

According to the model in this article, we can get the following conclusions:

(a) Trend of R&B and Country: The development trend between R&B and Country is relatively similar. It started in 1940, reached its peak in 1960-1980, and stabilized after 1980.

(b) Trend of Jazz and Vocal: The development trend of Jazz and Vocal is relatively similar. It began to rise in 1945, reached its peak in 1950-1965, and began to decline after 1970, indicating that people's attraction to it has declined.

(c) The indicator valence has the greatest im-pact on music popularity and is the most "contagious" indicator. The indicators energy, duration_ms, and explicit have a relatively large impact on music popularity and are also "infectious".

References

[1] Ma Qi. On the Development of Western Music in the 20th Century [J]. Popular literature and art, 2013(17): 142-143.
[2] Alaa Tharwat. Principal component analysis - a tutorial.2016, 3(3): 197-240.
[3] Liang Wei, Wang Duanxia, Zhang Li. Grey correlation analysis [J].Chinese Journal of Hearing and Language Rehabilitation Science, 2013(4): 292-295.