

# Music Style Analysis among Haydn, Mozart and Beethoven: an Unsupervised Machine Learning Approach

**Ru Wen**

Xi'an Jiaotong University  
wenru.1115@xjtu.edu.cn

**Zheng Xie**

Xi'an Jiaotong University  
xie-zheng123@stu.xjtu.edu.cn

**Kai Chen**

Xi'an Jiaotong University  
chenkai0208@stu.xjtu.edu.cn

**Ruoxuan Guo**

Xi'an Jiaotong University  
yuanfang123@stu.xjtu.edu.cn

**Kuan Xu**

Xi'an Jiaotong University  
xukuan@stu.xjtu.edu.cn

**Wenmin Huang**

Xi'an Jiaotong University  
wolfie@stu.xjtu.edu.cn

**Jiyuan Tian**

Xi'an Jiaotong University  
tt6688@stu.xjtu.edu.cn

**Jiang Wu**

Xi'an Jiaotong University  
jwu@sei.xjtu.edu.cn

## ABSTRACT

*Different musicians have quite different styles, which has influenced by their different historical backgrounds, personalities, and experiences. In this paper, we propose an approach to extract melody based features from sheet music, as well as an unsupervised clustering method for discovering music styles. Since that existing corpus is not sufficient for this research in terms of completeness or data format, a new corpus of Haydn, Mozart and Beethoven in MusicXML format is created for research. By applying this approach, similar and different styles are discovered. The analysis results conform to the Implication-Realization model, one of the most significant modern theories of melodic expectation, which confirms the validity of our approach.*

## 1. INTRODUCTION

Unique styles of musicians have been an attractive project for centuries. There are lots of characteristics used to recognize styles such as form, texture, harmony, melody, and rhythm [1]. Existing studies utilize those characteristics, as well as audio information, to extract feature for classification and retrieval tasks [2, 3, 4]. It's worth mentioning that melodic interval, which is measured as the distance in semitones between the two adjacent notes, carries strong information of music style [5]. According to current cognitive theories such as Implication-Realization model [6] and some tests on consecutive intervals [7], we could induce strong expectations on melodic continuations with only two consecutive intervals, a.k.a. a bigram. Such conclusion was used to identify accurately the transitions between the Baroque, Classical, Romantic, and Post-Romantic periods [4], or measure the evolution of contemporary Western Popular Music [1, 8].

As we all know, different musicians also have quite different styles, which has influenced by their different historical

backgrounds, personalities and experiences. A music work can express its composer's sentiment and character, so it's possible to figure out the composer based on his music work [9]. However, musicians in the same period, especially those in teacher-student relationships, may to some extent be influenced by each other. What's more, a composer's character may be influenced by life experiences and his music style may have some change during his lifetime. So this paper seeks to do some work in music feature extraction and style discovering and analyzing with the clustering methods.

Throughout the music history we select Haydn, Mozart, and Beethoven as research objects. Haydn was a friend and mentor of Mozart, and a teacher of Beethoven, so his music had great impact on the other two. He entered a Choir School when he was only five years old and was well musical enlightened, and began composing after he left the choir. His music was distinctive and boldly in individual, inspired by a form of heightened emotionalism known as 'Sturm and Drang'. Mozart was born in Archbishopric of Salzburg, which was a peaceful small town. He was employed as a court musician when he was in an early age but he chose to quit because of not receiving deserved esteem and treatment, so many of his early works are related to religious rites. Apart from these, Mozart was an adult child, so besides religious-themed music, most of his works are brisk and lively. Differing from Haydn and Mozart, Beethoven suffered from both political upheaval and physical diseases. He was greatly impacted by the thought of freedom, equality, and brotherhood, so he gradually composed with his own individual style and most works are grand and powerful [9].

We decided to use data in MusicXML format for the convenience of melody extraction. There are large music collections like Kunst der Fuge collection with about 18000 MIDI files (mostly piano works or reductions) contributed by the users. The collection in MusicXML format is much smaller, only contains 880 manually encoded compositions in 4116 movements [10], which cannot meet the need of completeness and accuracy of the three composers' work. Thus, we make our own database in the format of MusicXML.

In this context, our contribution is threefold. Firstly, we

Copyright: ©2016 Ru Wen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

build a database of Haydn, Mozart, and Beethoven, in MusicXML format, followed by melody extraction and feature extraction. Secondly, we propose an unsupervised machine learning approach by clustering the music into several clusters based on bigram probability distributions. Finally, together with some conclusions of Implication-Realization model, we give an explanation on the divisions and meanings of each cluster. Taking style inherit into consideration, we combine our conclusion with the development of music style and try to understand the continuity and evolution among the three musicians. The consistency of the unsupervised analysis results and the theory confirms the validity of our approach of feature extraction based on melody and unsupervised music style analysis.

## 2. PRELIMINARIES

### 2.1 Implication-Realization Model

In the middle of last century, melodic expectation was proposed in Meyer's *Emotion and Meaning in Music* [11], discussed the emotions and meaning with the perspective of cognition and expectation. In 1990, Narmour developed Implication-Realization model of melodic expectation based on Meyer's theory. This model is one of the most significant modern theories of melodic expectation. According to this theory, the perception of a melody continuously causes listeners to generate expectations of how the melody will continue.

In the theory of I-R model, *closure* states that the implication of an interval is inhibited when a melody changes in direction, or when a small interval is followed by a large interval. Other factors also determine closure, like metrical position (strong metrical positions contribute to closure), rhythm (notes with a long duration contribute to closure), and harmony (resolution of dissonance into consonance contributes to closure).

Under the above circumstance, when an interval doesn't form a closure, it will exert implications on how the listeners would expect the continuing melodies. The subsequent interval (formed by the next tone and the second tone of the first interval) is called realized interval. The realized interval may not conform to previous implications, while deviation from the implications often generate certain emotions and produce certain aesthetic influence. In the I-R model, considering that the implicative interval ranges from 0 to 11 semitones and that the realized interval is confined to one octave, the implicative interval is divided into large and small interval, six semitones being the threshold. And then five governing principles are presented, based on the melodic implications that defined by the tendency of pitch and the size of interval:

Registral direction: small intervals imply continuation of pitch direction; large intervals imply a change of direction.

Intervallic difference: small intervals imply similar-sized realized intervals; large implicative intervals imply smaller realized intervals.

Registral return: the second tone of a realized interval returns to the original pitch (within 2 semitones),

---

```

1: procedure K-MEANS( $X, k$ )  $\triangleright X = \{\mathbf{x}_{1\dots n}\}$ : dataset
2:   initialize  $C = \{\mathbf{c}_{1\dots k}\}$  at random  $\triangleright k$ : # clusters
3:   repeat
4:      $G_{1\dots k} \leftarrow \emptyset$ 
5:     for  $\forall \mathbf{x} \in X$  do
6:        $i \leftarrow \operatorname{argmin} \operatorname{distance}(\mathbf{x}, \mathbf{c}_i)$ 
7:        $G_i \leftarrow G_i \cup \{\mathbf{x}\}$   $\triangleright G_i$ : the  $i$ -th cluster
8:     end for
9:      $\mathbf{c}_i \leftarrow \frac{1}{|G_i|} \sum_{\mathbf{x} \in G_i} \mathbf{x}$ 
10:  until no centroid moved
11:  return  $C, G_{1\dots k}$ 
12: end procedure

```

---

Figure 1: K-means Algorithm

which forms symmetric (/aba/) or near-symmetric (/aba'/) patterns.

Proximity: realized intervals are often small, within 5 semitones.

Closure: implicative and realized intervals are opposite; besides, realized interval is smaller than implicative interval.

According to theories of I-R model, we have enough reason to believe that by extracting features from the two subsequent intervals formed by implicative and realized intervals, the style and emotions of the music can be demonstrated to a certain extent.

### 2.2 K-means Algorithm for Clustering

Clustering is a common machine learning task, which aims to grouping a set of instances into several clusters, and instances should be similar to those in the same cluster and different with those in different clusters. There are different clustering algorithms focuses on different criterion, e.g., connectivity-based clustering, centroid-based clustering, density-based clustering, etc.

K-means algorithm is one of the most common centroid-based clustering algorithm, it uses Euclidean distance as its similarity measure:

$$\operatorname{distance}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}, \quad (1)$$

where  $d$  is the dimension of the feature vector  $x$ . The algorithm optimizes the positions of cluster centers and the cluster assignment of each instance iteratively, The procedure is described in Figure 1.

## 3. COLLECTION OF MUSIC DATA

The databases existing on the Internet vary enormously not only in size and format, but also in quality and accuracy. Under this circumstance, collecting music data for computing tends to be a tough and time-consuming mission.

### 3.1 Format of Music Data

No matter what data chosen in a current database or downloaded on the Internet are collected, the format of data is

Haydn	Mozart	Beethoven
Piano Sonatas	Piano Duets	Trios for Piano
Masses	Piano Sonatas	Piano Sonatas
Piano Pieces	Piano Variations	Piano Duet / Quartet
Trios for Piano		Concerto for Piano
Piano Duos		Bagatelle for Piano
		Rondo for Piano
		Fantasia for Piano

Table 1: Pieces types of collected music scores.

always the first thing to be determined. As for music data, abundant formats are flooding in front of our eyes, such as MusicXML, MIDI, PDF, Sibelius, capella.

Considering the accuracy of music data and the completeness of music information, eventually we choose MusicXML as the main formats for further study.

MusicXML (Music Extensible Markup Language) [12], which is a standard open format for exchanging digital sheet music, is regarded as one the best formats that consists almost all the information in one piece of music we may utilize when computing.

Although MusicXML is the best choice for analyzing music data, numerous pieces of music are digitized only in MIDI format instead of XML. To systematically analyze the information contained in the score, all the MIDI files are converted into XML and some of the damaged files are replaced or restored after manual review. Then we check all the collected files and classify them by different music form via aural measure and divided them into several packages.

In summary, all the pieces of music scores we collected are listed in Table 1.

As is shown in Table 1, almost all the music scores related to piano that are composed by three musicians are concluded in our database. There might be some damaged or lost pieces of music on account of some historical reasons, however, the amount is large enough for further study on the styles of their piano composition as the representatives of classic music.

### 3.2 Establishment of Database

After deciding the music format that we shall utilize, then, it is time to establish the database for our study. Two options are available for obtaining suitable music data.

On the one hand, some existing databases can be used directly, such as Center for Computer Assisted Research in the Humanities at Stanford University, which is not only an excellent collection of classic music in MusicXML, but also a complete study done by this department in reference. Nevertheless, the total number is small (about 880), especially lack of the piano works by Beethoven. Some other collections also do help to our database, as an illustration, Kunst der Fuge collection and a collection from peach note.

On the other hand, as for some pieces of music which are not famous enough to be collected in current databases, we downloaded them on some specific websites, such as musescore<sup>1</sup> and musicalion<sup>2</sup>.

After gathering all pieces of music data we may use, then the attention is tended to the preprocessing of these massive

<sup>1</sup> <https://musescore.com/sheetmusic>

<sup>2</sup> <http://www.musicalion.com>

data. According to Implication-Realization model, numerous music patterns are hidden in the melody of the music, more specifically the melodic interval pattern. Therefore, we extract music melody from each piece of music and list them chronologically into our final database, which will be discussed in detail in the next section.

## 4. METHOD

### 4.1 Melody Extraction

Previous study shows that melodic interval pattern strongly indicates the various music styles in different music [4]. To analyze melodic interval pattern, melody should be extracted first. Salamon et al. proposed an approach of melody extraction from wave audio [13], but there is not any mature approach to extract melody from sheet music. In this situation, we extract music melody from each music score by choosing the highest note in each chord, and put them in a list in chronological order. To ease processing, each note was converted to an integer in terms of the number of semitones from the lowest note. Thus, each two integers in the list form a melodic interval.

### 4.2 Feature, Normalization and Flatten

Consider each pair of adjacent melodic interval no more than an octave, counting the frequency of occurrence of different pairs, we can code the frequency of occurrence into a 25 by 25 matrix  $M$ . Formally:

$$M^s = (m_{ij}^s) \in \mathbb{N}, \quad (2)$$

where  $m_{ij}^s$  is the frequency of occurrence of melodic interval pair  $(i - 12, j - 12)$  in the music score  $s$ . The Figure 2 gives a demonstration of the visualization of the matrices.

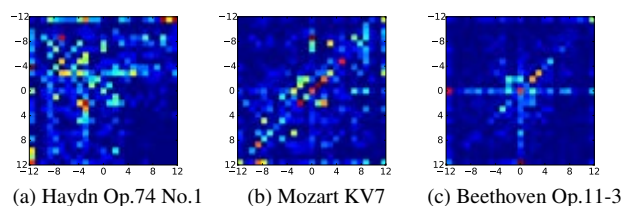


Figure 2: Demonstration of melodic interval pair matrix.

Because of the variation of the score length, data density varies significantly. To avoid negative effects, normalization is necessary. Consider two normalization method: Joint probability distribution  $P(i_1, i_2)$  and conditional probability distribution  $P(i_2|i_1)$ . The former is computed by normalizing  $M$  by its total sum while the latter is computed by dividing each element by the sum of corresponding row. Formally:

$$P_s(i_1, i_2) = \frac{m_{i_1, i_2}^{(s)}}{\sum_{i, j} m_{i, j}^{(s)}}, \quad (3)$$

$$P_s(i_2|i_1) = \frac{m_{i_1, i_2}^{(s)}}{\sum_j m_{i_1, j}^{(s)}}. \quad (4)$$

In order to use distribution matrices as input for clustering, it's necessary to flatten the matrices into vectors:

$$\mathbf{x}^{(s)} = (a_0^{(s)}, \dots, a_{624}^{(s)}, b_0^{(s)}, \dots, b_{624}^{(s)}), \quad (5)$$

where

$$a_{i \times 25 + j}^{(s)} = P_s(i, j) \times 25, \quad (6)$$

$$b_{i \times 25 + j}^{(s)} = P_s(i|j). \quad (7)$$

For purpose of combining the information that joint and condition probable distribution, we code two matrices into one vector. It is notable that the elements of joint probable distribution matrices should be multiple by 25, the number of rows of the matrices, to ensure two distribution vectors are in the same scale.

### 4.3 Clustering

With the processed feature vectors, K-means algorithm could be applied. To determine the number of clusters, we used *elbow point method*. We plot how that the clustering cost changes as a function of the number of clusters. If the cost decreases sharply at some point but slows down its decrease after that, this point should be a good number of clusters. The cost function is defined as Equation 8:

$$Cost = \log\left(\frac{1}{n} \sum_{i=1}^k \sum_{\mathbf{x} \in Cluster_i} |\mathbf{x} - Centroid_i|^2\right) \quad (8)$$

According to the curve shown in Figure 3, we chose 4 the number of clusters.

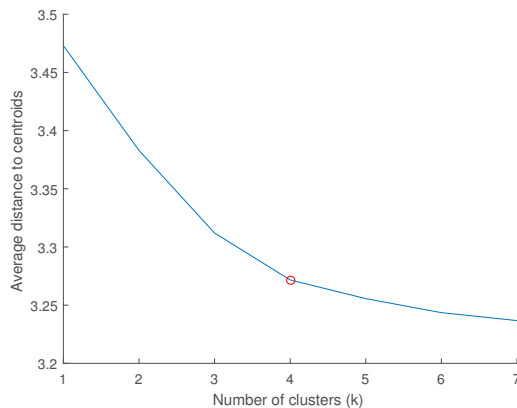


Figure 3: Cost variation with time. Elbow point occurs at 4.

The traditional K-means uses a group of random seeds as the initial cluster centers, which may lead to a bad clustering results when the initial centers differ from the real distribution of the data. We use K-means++ algorithm [14] to generate the initial cluster centers. K-means++ algorithm choose a group of initial cluster center from the instances by maximizing the distances of the centers in a randomized greedy manner.

### 4.4 Analysis Process

The total analysis process is shown in Figure 4.

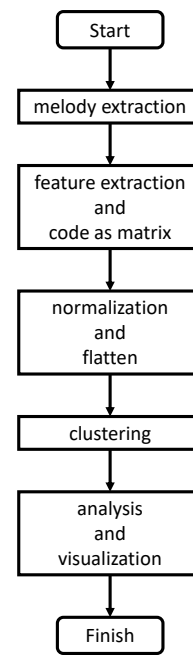


Figure 4: The flow chart of the analysis process.

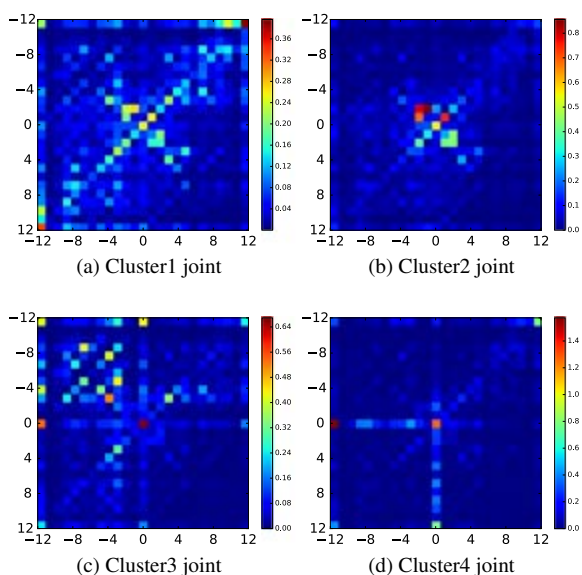
## 5. RESULTS

### 5.1 Clustering Results

We used K-means algorithm to cluster the pieces, finding out that 4 is a reasonable number of clustering, while the differences between cluster centers are getting inconspicuous when we cluster the pieces into more than 4 clusters. This conclusion is also identical to the result of elbow point method. As shown in Figure 5, four cluster centers differs from each other significantly, as well as some distinct patterns in them. The percentages of four kinds in three musicians' pieces are shown in Figure 6, and each percentage indicates the ratio of pieces classified into each cluster to all the pieces of a particular musician. Table 2 gives some example of the clustering results.

Music	Cls
<b>Haydn</b>	
Sonata No. 42 in D major, Hob. XVI:27	1
Sonata No. 4 in G major, Hob. XVI:G1	2
Presto in C major for Flute Duet, Hob. XIX:24	1
String Quartet No. 57 in C major, Op. 74 No.1, M4	1
Music Box for Harp	2
<b>Mozart</b>	
Piano Variations, K.24	2
Le nozze di Figaro, K.492	1
Piano Sonata No.11 in A major, K.331, M3	2
Piano Sonata No.17 in B-flat major, K.570, M1/3	2
Piano Sonata No.17 in B-flat major, K.570, M2	1
<b>Beethoven</b>	
Symphony No. 5 in C minor Fate, Op. 67, M1	4
Symphony No. 5 in C minor Fate, Op. 67, M2/3/4	3
Symphony No. 6 in F major Pastoral, Op. 68, M1/2/5	3
Symphony No. 6 in F major Pastoral, Op. 68, M3/4	4

Table 2: Some of the cluster results.



**Figure 5:** Visualization of four cluster centers (The shown matrices are joint conditional distribution part).

In addition, we used some internal clustering evaluation metrics, i.e., Davies-Bouldin index, Dunn index, and Silhouette coefficient, to evaluate the results (c.f. Table 3). We compared the metrics among many clustering algorithms, and confirmed that K-means is the best performed algorithm with the data used in this paper. Since the unsupervised internal evaluation metrics are not golden standard of the clustering results, the next subsection shows the results are meaningful in a music theoretical way.

Davies-Bouldin Index	Dunn Index	Silhouette Index
2.847	0.413	0.046

**Table 3:** Internal evaluation metrics for clustering.

## 5.2 Analysis with I-R Model

There are considerable proportions of pieces from all three musicians were classified into Cluster1, which suggests that Cluster1 is a common style at the time. Meanwhile, the rest of Haydn's pieces were classified into Cluster2 and Cluster3, but no Cluster4. Mozart has more pieces in Cluster2, and pieces in Cluster4 begin to appear. The latest one of the three, Beethoven, carried the pieces forward which are in the style implied by Cluster4, as well as kept the other styles.

Considerable proportions in all three musicians' pieces are clustered into Cluster1, implying that Cluster1 represents a common pattern. An obvious pattern in Cluster1 is shown in Figure 7a, which conforms to the Registral Return principle in Implication-Realization model. Registral Return principle describes the symmetric or near-symmetric melodic archetype like /aba/ or /aba'/ where the second tone of a realized interval is very similar to the original pitch (within 2 semitones) [6]. The pattern from Cluster1 embodied the melodic expectation of I-R model, evidenced the melodic progress of musicians' pieces meets the psychological expectation of people.

As shown in Figure 5b, cluster center of Cluster2 shows a pattern as Figure 7b. This pattern represents playing three adjacent tone in diatonic scale continuously, and is in conformity with the Intervallic Difference and Proximity principle. Melody with this characteristic often goes more gently, and expresses relaxing and pleasant mood, which is correspond to Mozart's music style. Pieces in Cluster2 mainly belong to Mozart, shows the characteristic of Mozart's.

In Figure 5c which is the cluster center of Cluster3, we can find that the bigrams mostly distributed in the top-left area, which means plenty of two (or more) successive downward melodic intervals and little upward intervals. This pattern does not correspond to any principles in I-R model, i.e., this kind of music are likely to be sorrowful or philosophical, and would inflame the strong feelings of the audiences by breaking the music expectation of them. This is distinct feature of Haydn's and Beethoven's. On the contrary, Mozart's pieces of music are more often positive and optimistic, so he has less percentage pieces classified to Cluster3.

We can also observe a pattern shown as Figure 7c in cluster center (0, 0) of Cluster4 (c.f. Figure 5d), which represents playing a same tone for three or more times. Besides, there are also two peaks at (0, -12) and (12, 0) in the visualization of cluster center of Cluster4, means that it's common to play a tone an octave higher or lower after of before two same tone played. Pieces with this feature appear primarily in Beethoven's, expressing a dramatic and belligerence music style. It's worth noting that we can also find the peaks, although gentle, in the Cluster3 (c.f. Figure 5c). This tells us that both styles in Cluster3 and Cluster4 are impactive and belligerent.

In summary, we discovered four music styles in the pieces of Haydn, Mozart, and Beethoven by a clustering approach. By visualize the cluster centers, we found that those styles conform the principles in I-R model to a large extent, and agree with the personalities of the three. This gives us a explanation of the clustering results in a music theoretical way, as well as a proof of the validity of our approach.

## 6. CONCLUSION AND FUTURE WORK

When we speak of the Classical period in music, the names of these three composers always come to mind: Haydn, Mozart, and Beethoven. The paths of these three great masters somehow crossed when they travelled to Vienna, which fused the music styles of them. On the other hand, they also have different characteristics in their music composition.

In this paper, we firstly establish a corpus of the three, and propose a melodic bigram feature extraction method. Further, we propose an unsupervised clustering method for discovering music styles without label. Our analysis results meet the I-R model of music expectation theory, which proves the effectiveness of the method, and recovers a set of factors that identifies different music styles among Haydn, Mozart, and Beethoven.

There are some possible ways to extent our work. Enlarge the range of the composers may improve the number of styles found, as well as the accuracy of the patterns. Further, we used an easy way to extract melody from the music, which can be improved by using an approach that takes tonic, dominant, and subdominant into consideration. In the last, the patterns we found are not mutually exclusive,

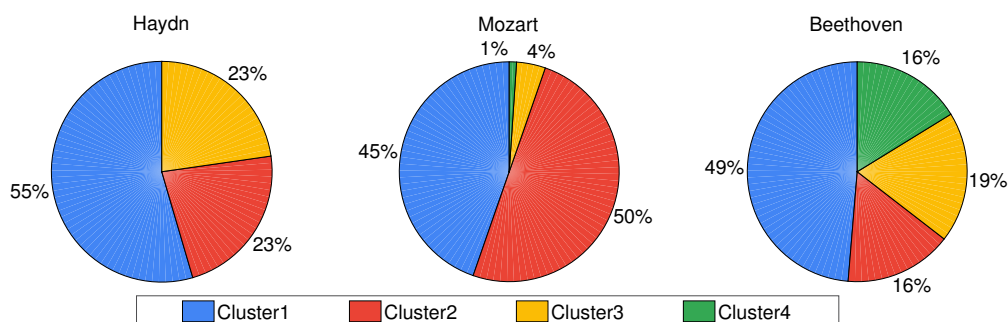


Figure 6: The percentage of pieces that fall into each cluster.

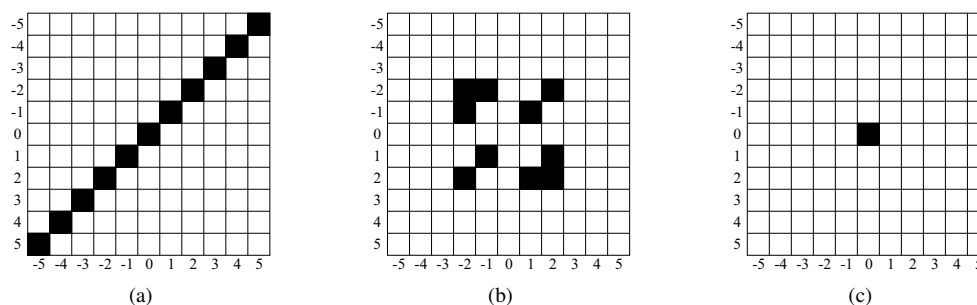


Figure 7: Patterns that found in visualization.

each cluster may contain several characteristics. Some dimension deduction or decomposition algorithms may be applied to the result matrices to separate the distinct music style characteristics.

## 7. REFERENCES

- [1] J. Serrà, Á. Corral, M. Bogoñá, M. Haro, and J. L. Arcos, "Measuring the Evolution of Contemporary Western Popular Music," *Scientific Reports*, vol. 2, jul 2012.
- [2] C. McKay and I. Fujinaga, "JSymbolic: A feature extractor for MIDI files," in *International Computer Music Conference, ICMC 2006*, New Orleans, LA, United states, 2006, pp. 302 – 305.
- [3] M. S. Cuthbert, C. Ariza, and L. Friedland, "Feature extraction and machine learning on symbolic music using the music21 toolkit," in *Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011*, Miami, FL, United states, 2011, pp. 387 – 392.
- [4] P. H. R. Zivic, F. Shifres, and G. A. Cecchi, "Perceptual basis of evolving Western musical styles," *Proceedings of the National Academy of Sciences*, vol. 110, no. 24, pp. 10034–10038, may 2013.
- [5] G. Vulliamy, N. A. Josephs, G. Holt, and D. Horn, "The New Grove Dictionary of Music and Musicians," *Popular Music*, vol. 2, pp. 245–258, 1982.
- [6] R. O. Gjerdingen and E. Narmour, "The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model," *Notes*, vol. 49, no. 2, p. 588, dec 1992.
- [7] E. Schellenberg, "Expectancy in melody: tests of the implication-realization model," *Cognition*, vol. 58, no. 1, pp. 75–125, jan 1996.
- [8] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, "The evolution of popular music: USA 1960-2010," *Royal Society Open Science*, vol. 2, no. 5, pp. 150081–150081, may 2015.
- [9] D. Hertz, *Mozart, Haydn and Early Beethoven: 1781-1802*. W W NORTON & CO INC, 2008.
- [10] V. Viro, "Peachnote: Music Score Search and Analysis Platform," in *Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011, Miami, Florida, USA, October 24-28, 2011*, 2011, pp. 359–362.
- [11] L. B. Meyer, *Emotion and Meaning in Music*. Chicago, Illinois, USA: The University of Chicago Press, 1961.
- [12] M. Good, "MusicXML for notation and analysis," *The virtual score: representation, retrieval, restoration*, vol. 12, pp. 113–124, 2001.
- [13] J. Salamon, E. Gomez, D. P. W. Ellis, and G. Richard, "Melody Extraction from Polyphonic Music Signals: Approaches, applications, and challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 118–134, March 2014.
- [14] D. Arthur and S. Vassilvitskii, "K-means++: The Advantages of Careful Seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.