# Distinguishing Chinese Guqin and Western Baroque pieces based on statistical model analysis of melodies

Yusong Wu[1] and Shengchen Li[1]

Beijing University of Post and Telecommunications, Beijing 100876, P. R. China
wuyusong@bupt.edu.cn
shengchen.li@bupt.edu.cn

**Abstract.** This paper proposes a method to determine different genres of melody according to the melodic interval of the melody with Western Baroque and Chinese Guqin music used as an example. Melodic interval histogram and Markov chains are proposed to differentiate Western Baroque and Chinese Guqin music, where the similarity is measured with Kullback-Leibler divergence. A significance test is done and the results show that our method is capable of distinguishing between Western Baroque and Chinese Guqin pieces. This conclusion further supports that extracting melodic interval features could be a possible way to distinguish symbolic music melodies from different genres.

**Keywords:** statistics, computational musicology, machine learning

## 1 Introduction

With the introduction of deep learning algorithms, automatic music composition has vastly developed in recent years. In the automatic music composition algorithms, a large proportion of the works are trained on a single music genre hence the generated music has the same genre as the music in training dataset. For instance, Bachbot [1] and DeepBach [2] composed polyphonic music in the style of Bach's Chorales; MidiNet [3] generated pop music melodies; Eck and Schmidhuber [4] proposed a system generating blues melody. Besides, there are models and systems that generate music in different genres. The WaveNet [5] is capable of generating music in different genres given conditional input, and the VRASH [6] system could generate melody in various genres with heuristic filtering.

Thus, for a certain generative model using a single set of parameters to generate music in multiple genres, it is necessary to differentiate and evaluate the music genre of the music generated for the validation of the generative model. The purpose of this paper is to propose a method that could determine different music genres, where Bach Chorales music and Chinese Guqin music are used as an example.

It is herein proposed to use a computational method employing interval histogram and Markov chains to distinguish Western Baroque pieces and Chinese

Guqin pieces. Melodies are analyzed in melodic interval sequence. A five-Fold cross-validation is applied and the dataset is split into training set and test set. Firstly, a melodic interval histogram is drawn and a Markov chain is trained for the whole training set and on each music piece in test set. Then, for each music piece in test set, similarity using Kullback-Leibler divergence is measured between that music piece to each of the two genres. Last, to verify the effectiveness of our method, a paired t-test is used to verify the significant difference between the two similarities measured for the two genres. The result of the paired t-test shows that both melodic interval histogram and Markov chain are capable of distinguishing Western Baroque pieces and Chinese Guqin pieces.

The remainder of the paper is organized as follows. Related works are first presented in section 2. Dataset and data representation used in this paper are described in section 3. Statistic models, algorithms are introduced in section 4. In section 5, experiment setups and experimental results are presented. Section 6 includes our explanation of the results. Conclusions are drawn and some future work is proposed in section 7.

## 2 Related Works

In this paper, melodies are represented as melodic intervals. Melodic interval histogram and Markov chain are used to represent the melodic interval features of the melody. Melodic interval refers to the distance between the two pitches of the two notes when played in sequence [7]. The melodic interval expresses the temporal sequence in terms of the distance between two adjacent pitches [8].

Representing symbolic music data in melodic intervals often achieves good performance and is widely used in music genre recognition (MGR) tasks. According to Correa [8], melodic interval representation offers more discriminative information than using absolute pitches or pitch contours. In [9], De Leon and Quereda compare the ability to separate two particular musical styles among a number of melodic, harmonic, and rhythmic statistical descriptors. Using the melodic interval feature, De Leon and Quereda achieved 100% test accuracy, and further conclude that melodic interval feature is one of the most discriminate features. In [10], Chai and Vercoe compared several representations of melodies when classifying European folk music. The representations used in Chai and Vercoe's work are absolute pitch representation, absolute pitch with duration representation, melodic interval representation, and contour-based representation (contour-based representation quantizes melodic interval representation into five levels, indicating conjunct motion, small ascending or descending in adjacent note, and large ascending or descending in adjacent note). By comparing the accuracy achieved by different data representations, Chai and Vercoe found that the interval representation outperforms the absolute pitch representation and the contour-based representation. Besides, representing data as melodic interval also is invariant under pitch transpositions [8], thus giving the advantage of analyzing melodies regardless of the key. Therefore, in this paper, we use melodic interval for data representation, for not only the promising performance of the melodic interval obtained in previous works but also the merit of analyzing melody regardless of the pitch transpositions of the melody.

In addition to melodic interval representation, melodic interval histograms reflect the information about the order in note sequence and relative frequency of the intervals. Knopoff [11] used melodic interval histogram to analyze the melodic activity in Bach's Fugue collection. Simsekli [12] uses a melodic interval histogram to classify music genre using bass lines on a 3-root 9-leaf label MIDI dataset. As much as the accuracy of 100.00% for the root labels and 84.44% for the leaf labels were achieved, showing the effectiveness of using melodic interval histograms to extract music features.

Furthermore, Markov chains are also good at capturing temporal patterns in music. Verbeurgt [13] used Markov chain to effectively extract patterns in music for composition.

## 3 Dataset and Data Representation

### 3.1 Dataset

The Western Baroque music and Chinese Guqin music are very unique music genres and have the following characteristics that facilitate in distinguishing these two music genres: 1) The Western Baroque music and Chinese Guqin music have significant differences in music style which can be highly distinguishable by people even with little or no music education backgrounds; and 2) In both Baroque music and Chinese Guqin music, most melodies do not contain overlapping notes, in which case every note or chord ends before another chord or note starts and if multiple tones are played at the same time, they must start and end at the same time. This feature greatly facilitates the modeling and analysis for we can now form the melody as a sequence of notes and chords.

Thus, Western Baroque music and Chinese Guqin music have been chosen in this paper as examples of two different genres. Specifically, we use Bach Chorales dataset[1] and a self-collected Chinese Guqin dataset[2] as our datasets.

Bach Chorales dataset is a symbolic music dataset formatted in MuiscXML, containing 409 pieces with 7241 measures. Bach Chorales dataset is mostly in four parts harmony, composed of four parts (Soprano, Alto, Tenor, Bass) in each score. In Bach Chorales dataset, some scores have extra accompaniment parts in addition to four chorales parts, such as violin, trumpet or timpani. The accompaniment parts often have few notes, with incomplete melodies and a large proportion of breaks. Because of the sparseness of the data in accompaniment parts, we only preserved four chorales parts, and all the accompaniment parts are ignored. Each of the four vocal parts is considered as a separate music piece.

The Chinese Guqin dataset is collected by ourselves, based on several books of score collection published by Chinese Guqin professionals. The scores are

[1] Bach Chorales dataset: `https://github.com/cuthbertLab/music21/tree/master/music21/corpus/bach`

[2] Chinese Guqin dataset: `https://github.com/lukewys/Guqin-Dataset`

mainly written in numbered notation, along with the Chinese Guqin notations[3]. A self-designed transcript system is used for typing the numbered notation into text format and convert the text format into MusicXML files. It is worth noting that, although the Chinese Guqin scores contain other music notations such as ornaments, fingerings, and expression notation, such information are ignored. Only the pitch and duration of melodies in Chinese Guqin scores are obtained. The Guqin music data we collected are mostly monophonic, contains only about 5% of notes as chord. In order to compute melodic interval among chords, we count all the possible melodic interval transiting from pitches in one chord to the next chord or note. For scores containing multiple phrases, we regard each phrase as a different music piece. The Chinese Guqin dataset contains 247 pieces with 6107 measures in total.

### 3.2 Data Representation

Although both the melodic pattern and the rhythmic pattern are different between Western classical music and Chinese classical music, the performance of Chinese Guqin music is very expressive, such that the duration of the note would vary with different Guqin players and play styles. We therefore focus primarily on extracting characteristics of tonal feature in Bach Chorales music and Chinese Guqin music, regardless of the rhythm pattern in melody.

The Western music system is based on heptatonic scale including seven levels per octave in scale, while the Chinese music system is mostly in pentatonic scale including five levels per octave, derived from the cycle-of-fifths theory. There are different mode variants (such as major or minor) in both Western music and Chinese music. Thus, by using a melodic interval representation which is invariant under transpositions, the melody could be represented regardless of their key or mode.

In this paper, the melodies are represented in melodic interval sequence. The melodic interval is defined as the absolute value of pitch difference in semitone. The melodic intervals are denoted in 13 classes, indicating semitones of intervals range in [0,12]. Intervals that are not multiple of 12 would be modulo by 12, following the method used in multiple works [10] [11]. However, unlike previous works using positive and negative intervals to denote upward and downward of the melodies, we consider only the absolute value of pitch difference, for reducing the dimension of the distribution subsequently extracted. Also, for intervals that are multiples of 12 and not zero, they are indicated as 12, to capture the pure octave leap. Only melodic interval of 0 semitones would be recorded as 0, to capture the conjunct motion, i.e. repeated notes, in melody progression.

Given two adjacent pitches $(p_t, p_{t+1})$ in melody, where $p_t$ represents pitch at time $t$, we define the melodic interval $i_t$ at time $t$ and the function of computing the interval $\mathrm{INT}(p_t, p_{t+1})$ as follows:

[3] The introduction of Chinese numbered notation Chinese Guqin notations can be found at `https://en.wikipedia.org/wiki/Numbered_musical_notation` and `https://en.wikipedia.org/wiki/Guqin_notation`

$$\mathrm{INT}(p_1, p_2) = i_t = \begin{cases} |p_{t+1} - p_t| \bmod 12 & \mathbf{if} |p_{t+1} - p_t| \nmid 12 \\ 0 & \mathbf{if} |p_{t+1} - p_t| = 0 \\ 12 & \mathbf{otherwise} \end{cases} \tag{1}$$

As we mentioned above, all the melodies in our dataset can be written as a sequence of notes and chords. Since we only care about the melodic progression, i.e. pitch progression, we can represent notes and chords as pitch set, regardless of the duration of notes and chords. A note, which includes only one pitch, would correspond to a pitch set containing only one pitch; a chord, which corresponds to multiple pitches, would correspond to a pitch set containing multiple pitches. We denote the pitch set at time $t$ as $P_t$. Thus, we can form the sequence of notes and chords into sequence of pitch sets $\{P_1, P_2, P_3 \dots P_m\}$.

Given a pair of pitch sets containing two adjacent pitch $(P_t, P_{t+1})$, the interval set for such pitch pair $I_t$ is defined as the set of all possible combination of progression from one pitch in $P_t$ to next pitch $P_{t+1}$, considering that all the transitions in note pitch contains indispensable information of melodic progression.

In other words, given a pitch set $(P_t, P_{t+1})$, the interval set of the pitch set would be:

$$I_t = \{\mathrm{INT}(p_i, p_j)\} \quad \forall p_i \in P_t \ , \ \forall p_j \in P_{t+1} \tag{2}$$

One example of our data representation is shown in Fig. 1. The pitch set representation of the example, in MIDI absolute pitch, is $\{[60], [57], [57, 45], [64], [62], [60], [62], [64, 52]\}$, in which each square bracket represents a pitch set and each number in square bracket represents a pitch. The melodic interval set is calculated as $\{[\mathrm{INT}(60, 57)]$, $[\mathrm{INT}(57, 57), \mathrm{INT}(57, 45)]$, $[\mathrm{INT}(57, 64), \mathrm{INT}(45, 64)]$, $[\mathrm{INT}(64, 62)]$, $[\mathrm{INT}(62, 60)]$, $[\mathrm{INT}(60, 62)]$, $[\mathrm{INT}(62, 64), \mathrm{INT}(62, 52)]\}$, with a result of $\{[3], [0, 12], [7, 7], [2], [2], [2], [2, 10]\}$, in which each square bracket represents an interval set and each number in square bracket represents the interval in the interval set. The value of the number denotes the interval value in semitone.

Fig. 1: One music segment example for data representation. The melodic intervals are computed using all possible combinations of the difference in pitch for adjacent notes. The melodic interval sequence for this music segment example would be: $\{[3], [0, 12], [7, 7], [2], [2], [2], [2, 10]\}$. Each square bracket represents an interval set and each number in square bracket represents the interval in the interval set. The value of the number denotes the interval value in semitone.

# 4 Method

## 4.1 Melodic Interval Histogram

We extract the melodic interval histogram as one feature. The distribution of melodic intervals are extracted in melodic interval histograms. By comparing the probability in different interval, we can examine how music in different genres are different in the use of melodic interval and melodic progression.

For each genre, the melodic intervals are counted on all the music pieces in that genre. The frequency of melodic interval is then normalized, presented as a histogram. For genre $G$, the melodic interval histogram IH is calculated as follows:

$$\mathrm{IH}^G\{x = i\} = \frac{\mathrm{count}(x = i)}{N_I} \tag{3}$$

where $N_I$ is the total number of melodic intervals counted and $\mathrm{count}(x = i)$ denotes the number counted for interval $i$, $i = 0, 1, 2 \ldots 12$.

## 4.2 Markov Chain

Although the melodic interval histogram contains statistic features of a music genre, the information in melodic interval histogram is limited in the static aspect. In other words, the melodic interval histogram only reveals the pattern in how melodic intervals are distributed but show no clue of how the melodic intervals are progressed. To extract the temporal pattern of the intervallic progression, we trained a Markov chain using the melody on each genre of music. By analyzing the transition matrixes in the Markov chains, we could extract patterns in melodic interval transition, i.e., the pattern of the melodic interval progression.

A Markov chain is a stochastic model describing the probability of the next state depends only on the current state. The probability of transferring from state $i$ to state $j$ is called transition probability $p_{ij}$, namely:

$$p_{ij} = \Pr\left(X_1 = j | X_0 = i\right) \tag{4}$$

We train the Markov model by counting the transition number and then normalize the transition count to transition probabilities. After we trained a Markov model, a parameter matrix is obtained containing all the transition probabilities. The parameter matrix is called the transition matrix of the Markov model. The transition matrix $P = (p_{ij})$.

In this paper, for each genre, a Markov chain is trained using all the music pieces in that genre. The state of Markov chain is the melodic interval, consisting of 13 states representing melodic intervals from 0 to 12, in semitone. In other words, the Markov chains is trained on melodic interval sequence. In our melodic interval sequence, multiple intervals occur in one timestep (as interval set). To resolve this problem, we count all the possible combinations of melodic interval transition, similar to how we define the interval set.

After a Markov chain is trained, the transition matrix $\mathbf{M}$ is obtained. We use $\mathbf{M}$ instead of $P$ to denote the transition matrix for distinguishing from the pitch set in the previous section. As a total of 13 melodic intervals are used in melodic interval representation in this paper, the dimension of the transition matrix is $\mathbf{M}$ is $13 \times 13$.

### 4.3 Similarity Evaluation using Kullback-Leibler Divergence

With the two probabilistic models being built on the melodies in two genres respectively, for each genre, two distributions are obtained, namely melodic histogram and transition matrix in Markov chain. In other words, $\mathrm{IH}^{Bach}$, $\mathbf{M}^{Bach}$, $\mathrm{IH}^{Guqin}$, and $\mathbf{M}^{Guqin}$ are obtained. The superscript on melodic interval histogram IH and transition matrix $\mathbf{M}$ denotes the source on which the distribution is calculated. In this way, the melodic interval pattern of given music genres can be represented in melodic histogram and transition matrix. Given a new piece of music melody $S$, we can extract the same two distributions on them, namely interval distribution $\mathrm{IH}^S$ and interval transition matrix for that music piece $\mathbf{M}^S$, and measure how similar the distributions of the given music piece are from the distributions extracted from Western Baroque pieces and Chinese Guqin pieces.

Kullback-Leibler divergence is an algorithm used to measure the similarity between two distributions. It is an intuitive and explainable method to use Kullback-Leibler Divergence to measure the similarity between the distribution of the feature.

Let $P(x)$ and $Q(x)$ be two probability distributions of a discrete random variable $x$. The Kullback-Leibler divergence between $P$ and $Q$ is defined as follows:

$$D_{\mathrm{KL}}(P \| Q) = \sum_{x \in \mathcal{X}} P(x) \log \left( \frac{P(x)}{Q(x)} \right) \tag{5}$$

The Kullback-Leibler divergence is always non-negative. The more similar two distributions are, the less Kullback-Leibler divergence value is computed. In other words, the lower the Kullback-Leibler divergence value, the higher the similarity between the two distributions. If two distributions are identical, the Kullback-Leibler divergence computed between the two identical distributions would be 0. The Kullback-Leibler divergence is not symmetric. The "$\|$" in the notation of Kullback-Leibler divergence denotes this asymmetric, as $D_{\mathrm{KL}}(P \| Q)$ measures the amount of information lost when $Q$ is used to approximate $P$.

In this paper, we define the similarity of a sample $S$ to a certain genre $G$ as the Kullback-Leibler divergence from the distribution of feature extracted from sample $D^S$ and the same kind of distribution of feature extracted from genre $D^G$. Here, distribution $D$ is melodic interval histogram or transition matrix, specifically, $D \in \mathrm{IH}, \mathbf{M}$. The similarity is computed as follows:

$$D_{\mathrm{KL}}(D^G \| D^S) = \sum_{x \in \mathcal{X}} D^G(x) \log \left( \frac{D^G(x)}{D^S(x)} \right) \tag{6}$$

In practice, before computing the Kullback-Leibler divergence of two distributions, a small constant of $1e-5$ is added to each entry in both two distributions, to avoid subsequent calculation of Kullback-Leibler divergence on zero entries.

# 5 Experiments and Results

## 5.1 Cross-validation

A five-fold cross-validation is used to evaluate our algorithms. All the pieces are randomly shuffled and split into five subsets with approximate the same length of total measures. Training and testing are practiced five times. For each time, we use the one subset as test set and the other four subsets as training set. For each genre, we build a melodic interval histogram and train a Markov chain using the music pieces in the training set. That is, we calculate $\mathrm{IH}^{Bach}, \mathbf{M}^{Bach}, \mathrm{IH}^{Guqin}$, $\mathbf{M}^{Guqin}$ on entire training set of each genre. For each music piece in test set, the interval histogram and Markov chain are built and $\mathrm{IH}^{S}$, $\mathbf{M}^{S}$ are obtained. Then, Kullback-Leibler divergence value of $D_{\mathrm{KL}}(\mathrm{IH}^{Bach}\|\mathrm{IH}^{S})$, $D_{\mathrm{KL}}(\mathrm{IH}^{Guqin}\|\mathrm{IH}^{S})$, $D_{\mathrm{KL}}(\mathbf{M}^{Bach}\|\mathbf{M}^{S})$, $D_{\mathrm{KL}}(\mathbf{M}^{Bach}\|\mathbf{M}^{S})$ are computed.

## 5.2 Significance Test using Paired t-test

If our method is able to extract distinguishable feature, given a sample from test set $S^G$ of genre $G$, the similarity of the sample to the genre $G$ should be significantly higher than the similarity to the other genre $G'$, $G' \ni G$. In our case, the Kullback-Leibler divergence value is inversely proportional to similarity. Thus, given a test sample $S^G$ from genre $G$, and distribution $D^S$ extracted from genre $S$, we should have:

$$D_{\mathrm{KL}}(D^G\|D^S) < D_{\mathrm{KL}}(D^{G'\ni G}\|D^S) \tag{7}$$

A paired t-test is applied to verify the significant differences between Kullback-Leibler divergence values on two genres, and a one-tail p-value is calculated to examine the significance of the hypothesis above. We choose the significant level of $p = 0.01$. That is, if the p-value calculated from the paired t-test is lower than 0.01, we can say for a given sample $S \in G$, it is statistically significant that $D_{\mathrm{KL}}(D^G\|D^S)$ is larger than $D_{\mathrm{KL}}(D^{G'\ni G}\|D^S)$.

The paired t-test is applied as follows. A set of paired data with $n$ pairs $X_i$ and $Y_i$ are given. In our case, of a single pair $X_i$ and $Y_i$ are the similarities measured from one distribution extracted from a music piece to each two distributions extracted from two music genre, namely $D_{\mathrm{KL}}(D^G\|D^S)$ and $D_{\mathrm{KL}}(D^{G'\ni G}\|D^S)$. First, the mean difference $\bar{d} = Y_i - X_i$ of the pairs is computed. Next, the standard deviation of the differences $s_d$ is calculated, and using $s_d$, standard error of the mean difference is calculated as $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$. Then, the t-statistic $t$ is calculated as $t = \frac{\bar{d}}{SE(d)}$. The t-statistic $t$ follows a t-distribution with degrees of freedom $df = n-1$, and comparing $t$ to the $t_{n-1}$ distribution, the p-value $p$ is obtained. The p-value $p$ , is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true.

### 5.3 Results

The melodic interval histogram of Bach Chorales collection and Chinese Guqin collection on the whole dataset are shown in Fig. 2. It can be observed that the major difference between Guqin collection and Bach Chorales collection is the use of the minor second and minor third. The Bach Chorales collection is abundant in the minor second while the Chinese Guqin music is more preferable to the minor third. Also, from the melodic interval histogram, we can observe that Bach Chorales collection is relatively extensive in the intervals of 0, 1 and 2 semitones, while having some proportion of intervals of 3, 4, 5, 7 and 12 semitones. In Chinese Guqin music, the proportion of the interval of 12 semitones is relatively high.

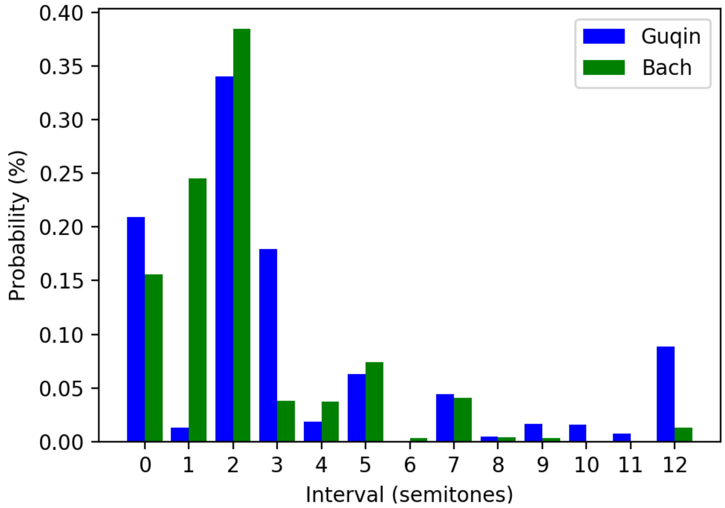
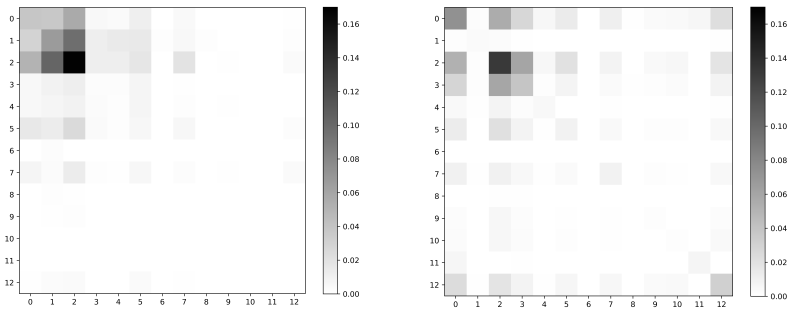



Fig. 2: The melodic interval histogram of Chinese Guqin collection (blue) and Bach Chorales collection (green).

The transition matrix in Markov models trained on the whole dataset of Chinese Guqin and Western Baroque pieces are shown in Fig. 3. It can be observed that the melodic interval transition distribution in Bach collection concentrates on the upper left of the matrix, while the interval transition distribution in Guqin collection is more separated. This reflects that the Bach collection has less large intervals, while the melodic progression in Chinese Guqin collection involves more big leaps in melodic intervals.



(a) Bach (b) Chinese Guqin

Fig. 3: The transition matrix of Bach Chorales collection (a) and Chinese Guqin collection (b). Numbers in the horizontal and vertical axes represent the interval value in semitone. The darkness of the element in the matrix denotes the probability of the transition.

The paired t-test results for KullbackLeibler divergence values are presented in Table. 1a and Table. 1b. Mean-difference($\bar{d}$), degrees-of-freedom($df$), statistic($t$) and the one-tail p-value($p$), of each of the five-fold cross-validation for the two genres, are presented. The results show that a p-value less than 0.01 is obtained in every significance test. This suggest that for a given sample in test set, $D_{\mathrm{KL}}(ID^G \| ID^S)$ is significantly lower than $D_{\mathrm{KL}}(ID^{G' \ni G} \| ID^S)$, and $D_{\mathrm{KL}}(\mathbf{M}^G \| \mathbf{M}^S)$ is significantly lower than $D_{\mathrm{KL}}(\mathbf{M}^{G' \ni G} \| \mathbf{M}^S)$. The results also show that transition matrix is better in distinguishing melodies, with larger mean-differences and lower p-values.

## 6 Discussion

Our results on melodic interval histogram of Bach Chorales music are similar to those of Knopoff and Hutchinson [11], who found that the conjunct movement is preponderance in Bach's fugue pieces, and movements by a major or minor second constitute about 70% of all pitch movements. Also, from results on melodic interval histogram, the lack of minor second in Chinese Guqin collection is mainly due to the musical scale in Chinese music pentatonic scale system not containing minor second. The abundant of interval of 12 semitones shows that Chinese Guqin music has a unique preference of using octave leap in melodic progression.

The sparseness differences of the transition matrix between Western Bach Chorales music and Chinese Guqin music confirm the intuitive fact that Bach collection often has more consecutive melody lines, whereas the Chinese Guqin collection is jumpy in melodic progression. This reflects the difference in the mode system between Chinese classical music and Western classical music. Furthermore, the jumpy nature in Chinese Guqin melody comes from octave harmonizations of the melody (as in the example of Fig. 1) and due to its common use in Guqin composing to raise or lower the notes by octaves in melody [14]. In-

(a) Melodic interval distribution

| | $\bar{d}$ | $df$ | $t$ | $p$ |
|---|---|---|---|---|
| Guqin 1 | -1.72 | 47 | -13.52 | 2.03e-18 |
| Guqin 2 | -1.88 | 48 | -20.44 | 6.27e-26 |
| Guqin 3 | -1.81 | 48 | -16.15 | 1.22e-21 |
| Guqin 4 | -1.67 | 49 | -14.43 | 7.10e-20 |
| Guqin 5 | -1.79 | 56 | -16.36 | 1.37e-23 |
| Bach 1 | 0.66 | 79 | 27.99 | 2.31e-43 |
| Bach 2 | 0.64 | 85 | 44.52 | 2.77e-61 |
| Bach 3 | 0.66 | 77 | 29.65 | 1.81e-44 |
| Bach 4 | 0.68 | 80 | 36.74 | 1.87e-52 |
| Bach 5 | 0.64 | 81 | 37.24 | 2.53e-53 |

(b) **Transition matrix**

| | $\bar{d}$ | $df$ | $t$ | $p$ |
|---|---|---|---|---|
| Guqin 1 | -1.84 | 47 | -13.25 | 4.32e-18 |
| Guqin 2 | -2.21 | 48 | -26.47 | 6.90e-31 |
| Guqin 3 | -2.1 | 48 | -17.8 | 2.22e-23 |
| Guqin 4 | -2.07 | 49 | -19.05 | 6.70e-25 |
| Guqin 5 | -2.22 | 56 | -21.19 | 5.13e-29 |
| Bach 1 | 1.98 | 79 | 51.39 | 4.42e-63 |
| Bach 2 | 2.06 | 85 | 54.67 | 1.31e-68 |
| Bach 3 | 2.03 | 77 | 48.79 | 2.88e-60 |
| Bach 4 | 2.06 | 80 | 69.22 | 9.27e-74 |
| Bach 5 | 1.95 | 81 | 51.99 | 1.38e-64 |

Table 1: The paired t-test result for Kullback-Leibler results of melodic interval distribution (a) and transition matrix (b). The mean-difference ($\bar{d}$), degrees-of-freedom ($df$), statistic ($t$) and the one-tail p-value ($p$) are presented. The number followed the genre represents the number of the five-fold cross-validation. The transition matrix is better in distinguishing melodies, as the table of transition matrix results in larger mean-difference and lower p-value.

terestingly, both two genres have maximum probabilities on the transition from the interval of second to the interval of the second.

From both Table. 1a and Table. 1b, we can see p-values much lower than 0.01 in each paired t-test. The low p-value supports that it is statistically significant that, given a music piece, the similarity measured to one genre is different from the similarity measured to another genre. This statistical conclusion suggests that both two distributions extracted from the two genres are significantly different from each other. Thus, both melodic interval histogram and Markov chain are capable of capturing unique characteristics in melodies in Chinese Guqin and Western Baroque pieces and distinguish between the two music genres. The mean difference in paired t-test of the transition matrix is larger than the mean difference in paired t-test of melodic interval histogram, suggesting that the feature extracted by Markov chain is more distinguishable or separable than melodic that the feature extracted by interval histogram. Therefore, Markov chain is better than interval histogram in differentiating Western Baroque pieces and Chinese Guqin pieces.

## 7 Conclusion

By building melodic interval histogram and Markov chains, features of melodic intervals are successfully extracted from melodies in Chinese Guqin and Western Baroque pieces. Kullback-Leibler divergence is used to compute the similarity of the given music piece to the music genre. A significance test has been carried out and the results show that our methods are capable of distinguishing Western Baroque and Chinese Guqin pieces. The Markov chains performs better in distinguishing Western Baroque and Chinese Guqin pieces. The success of our model further suggests that melodic interval features can serve as an effective way to extract characteristics of symbolic music melody.

## References


1. Feynman Liang. Bachbot: Automatic composition in the style of bach chorales. *University of Cambridge*, 8:19–48, 2016.
2. Gaëtan Hadjeres, François Pachet, and Frank Nielsen. Deepbach: a steerable model for bach chorales generation. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1362–1371. JMLR. org, 2017.
3. Li-Chia Yang, Szu-Yu Chou, and Yi-Hsuan Yang. Midinet: A convolutional generative adversarial network for symbolic-domain music generation. *Proceedings of the 18th International Society for Music Information Retrieval Conference ISMIR 2017*, 2017.
4. Douglas Eck and Juergen Schmidhuber. A first look at music composition using LSTM recurrent neural networks. Technical report, IDSIA USI-SUPSI Instituto Dalle Molle, 2002.
5. Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016.
6. Alexey Tikhonov and Ivan P Yamshchikov. Music generation with variational recurrent autoencoder supported by history. *arXiv preprint arXiv:1705.05458*, 2017.
7. Theta Music Trainer. Melodic intervals. `https://trainer.thetamusic.com/en/content/melodic-intervals`.
8. Débora C. Corrêa and Francisco Ap. Rodrigues. A survey on symbolic data-based music genre classification. *Expert Systems with Applications*, 60:190–210, oct 2016.
9. Pedro J Ponce De León and José Manuel Iñesta Quereda. Statistical description models for melody analysis and characterization. *International Computer Music Conference, ICMC 2004*, 2004.
10. Wei Chai and Barry Vercoe. Folk music classification using hidden Markov models. In *Proc. of the International Conference on Artificial Intelligence IC-AI'01*, volume 6. sn, 2001.
11. L. Knopoff and W. Hutchinson. An index of melodic activity. *Interface*, 7(4):205–229, dec 1978.
12. Umut Simsekli. Automatic music genre classification using bass lines. In *2010 20th International Conference on Pattern Recognition*, pages 4137–4140. IEEE, 2010.
13. Karsten Verbeurgt, Michael Dinolfo, and Mikhail Fayer. Extracting patterns in music for composition via Markov chains. In *Proceedings of international conference on industrial, engineering and other applications of applied intelligent systems*, pages 1123–1132. Springer, 2004.
14. Wang Z.Y. *Analysis of Guqin collection (In Chinese)*, chapter 4, page 143. Central Conservatory of Music Press, 2005.
15. Jean-Julien Aucouturier and Francois Pachet. Representing musical genre: A state of the art. *Journal of New Music Research*, 32(1):83–93, 2003.
16. Michael Scott Cuthbert and Christopher Ariza. music21: A toolkit for computer-aided musicology and symbolic music data. *Ismir*, 2010.