International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 921-928

https://doi.org/10.70517/ijhsa46485

# A Characteristic Analysis of National Tuning Scale Structures in Chinese Piano Music Composition

**Ye Yang**[1,*]
[1] Normal College, Xuzhou University of Technology, Xuzhou, Jiangsu, 464000, China
Corresponding authors: (e-mail: yangyexzit@163.com).

**Abstract** Chinese national music culture has a long history, and pentatonic modulation, as an important part of traditional music, plays a unique role in modern piano music composition. In this study, a national tuning recognition system was constructed based on the pitch level distribution matrix (PCDM) and temporal adaptive neural network (TANN), and Chinese piano music was spectrally analyzed using the constant Q transform (CQT). The study establishes a Chinese ethnic modal classification system covering 360 combinations by experimentally analyzing eight MIDI samples of Chinese folk songs, including four pentatonic modal and four heptatonic modal works. The experimental results show that the accuracy of the proposed algorithm in recognizing Chinese folk music modes reaches 100%, in which the pentatonic modal works such as “Jasmine Flower” are accurately recognized as G levied pentatonic modes, and the septatonic modal work “It's a Man of Our Knowledge”, with a 15.69% partiality, is successfully recognized as Yanle G levied septatonic modes. The study proves that the modal detection method based on PCDM features has significant advantages in the analysis of Chinese folk music, and provides an effective technical support for the digitization research of traditional music.

**Index Terms** Level distribution matrix, temporal adaptive neural network, constant Q transform, ethnic modal recognition, pentatonic modal, heptatonic modal

## I. Introduction

The world music culture is a pluralistic and intersecting cultural whole, and each era and each nation has its own cultural and historical background, and forms music forms with distinctive individual characteristics in this background [1]-[3]. And apart from the differences in history, folklore, language, religion and aesthetic habits of different musical forms, they are often reflected in the micro-variations of music [4], [5]. Among them, modulation is a form of musical tone organization structure summarized and created by human beings in long-term musical practice, and this form is often the basis for the development and change of musical styles and musical ideas [6]. Different ethnic groups and different historical periods, the formation of a hundred different, ever-changing modal categories [7]. The organization of modes and the interrelationships between modal levels, the direction of tones and the rhythmic laws of intervals, and the practice and differences of music theory are all important factors that constrain or influence musical styles [8]-[10]. Therefore, the study of tonality is important for mastering familiarity with national or local musical languages as well as for harmonic exploration [11].

Compared to traditional Chinese music, which is monophonic and develops melody through linear thinking, Western piano music is polyphonic and develops melody through three-dimensional thinking [12], [13]. How to base on the Chinese national tunes and handle the harmony between Chinese tunes and western harmonies in piano arrangements, which not only reflects the essence of piano composition but also reflects the Chinese national characteristics, is always a problem that people in the industry should think about and pay attention to [14]-[17].

In this study, a complete set of ethnic tuning analysis system is constructed by introducing advanced music information retrieval technology and machine learning algorithms. Firstly, the level distribution matrix technique is applied to extract two-dimensional features from the music signals, which can capture the level distribution and octave information at the same time, providing a rich data base for the subsequent tonal recognition. Then a temporal adaptive neural network is used to train the extracted features for deep learning, and a recognition model covering multiple tonal types such as pentatonic, hexatonic, and heptatonic is established. Through a large number of experimental validations on Chinese folk songs and piano works, the system not only accurately recognizes traditional pentatonic modes, but also effectively handles complex modal structures with the addition of partials, providing a reliable technical means for the quantitative analysis of the characteristics of ethnic modes in Chinese piano music composition.

# II. National scale

## *II. A. Style and tonality*

Style is the unity of content and form, the organic unity of subjectivity and objectivity. In the field of art, style always implies that some kind of iconic features show their own internal characteristics in external performance. In the field of music, the term style is more commonly used, and its connotation and extension are more complex. In music ontology, music style refers to the stable and systematic sensual characteristics of music in terms of structure, which involves many elements of music, such as melody, modulation, rhythm, harmony, timbre and so on. Among the various elements of music, modulation occupies a very important position. “The style, character, and color of music are first reflected in the melody, which is based on a certain key.” Modulation refers to a number of musical notes of different heights, around a sense of stability of the central tone, according to a certain intervallic relationship between the organization and the formation of the organic system.

## *II. B. The stylistic logic of Chinese pentatonic modes*

### II. B. 1) Chinese pentatonic modes

Pentatonic modulation is a widely used modulation in the world, but the pentatonic modulation used by different nationalities and regions is not the same. The Chinese pentatonic mode is a tonal system based on five tones: Gong, Shang, Horn, Zheng and Feather. Its formation has a long history, but it is difficult to say exactly when it came into being. During the Spring and Autumn Period, the calculation method of pentatonic modulation was recorded in “Guan Zi - Di Yu Chuan”, which is also the earliest formation of tonal calculation method in the world [18]. Chinese pentatonic modulation has evolved and developed in history, gradually forming pentatonic and hexatonic and septatonic modulations formed by adding partials to pentatonic, thus forming the “pentatonic” modulation system. Since China is a multi-ethnic country, due to its different languages, cultures and environments, the modes used in music are also more diverse, but in the music of Han Chinese and many ethnic minorities, the pentatonic mode occupies a dominant position and is widely used in Chinese traditional music. Chinese pentatonic modulation is the root and soul of Chinese music, representing the basic style of Chinese traditional music [19].

### II. B. 2) Scale fundamentals

The scale is the basis and basis for the formation of the tonality, and the Chinese pentatonic scale is formed on the basis of the pentatonic scale. The concept of "five tones" is closely related to the ideas of five elements, five things, and five bodies in traditional Chinese culture, and is the essence of traditional Chinese music culture. "Feng Wusheng" is to take Wusheng as the core and as the backbone. "Feng Wusheng" is the most common and important thinking method in traditional Chinese music, and it is also the basic morphological characteristic of forming the key of traditional Chinese music. In the traditional Chinese pentatonic system, the five positive tones can be used as the main tone of the key and thus organized into five keys, which are named after their main tones, namely the five keys of Gong, Shang, Jiao, Zheng and Yu. On the basis of the pentatonic mode, the six-tone and seven-tone tones formed by adding clear angles, changing signs, changing palaces, intercalations, etc., are also based on the five positive tones and the core, that is, "Feng five tones". Although traditional Chinese music, like Western art music, also uses the seven-tone scale, which can be said to be based on the natural seven tones, how to use the seven tones and what kind of thinking method to adopt are completely different. Western art music has an average of seven tones, while traditional Chinese music is "Feng Wusheng", that is, seven tones with five tones as the backbone. It is the tonal system dominated by the unique way of thinking of "Feng Wusheng" that has become an essential feature of traditional Chinese music, and "Feng Wusheng" has also become the basic style of Chinese pentatonic tonality.

### II. B. 3) Interval structure

Because of the pentatonic nature of the scale's foundation, the intervals within the modal scale are relatively unique. The intervals between the tones of the Chinese pentatonic scale include: major two, minor three, major three, pure four, pure five, minor six, major six, minor seven. The most distinctive feature of the intervals of the Chinese pentatonic scale is the lack of intervals with sharp tendencies such as semitones and triple perfect tones, which makes the whole scale seem “gentle”. Secondly, among the five intervals, there is only one major third (or minor sixth) formed by the intervals between the palace corners, which is also the only major third interval. The major third is the least harmonic interval in the pentatonic scale, while in Western music the major third is the most important core interval of harmony. The less use of major third intervals and the lack of sharp intervals reflect the “gentle” style of traditional Chinese pentatonic modes, which has a profound relationship with traditional Chinese aesthetic ideas and the characteristics of the Chinese nation. This “gentle” style of intervals has a profound relationship with traditional Chinese aesthetic ideas and the characteristics of the Chinese nation.

## III. Key technologies for ethnic modal analysis

### *III. A. Level Distribution Matrix (PCDM)*

The pitch level distribution matrix (PCDM) is a feature that describes the distribution of pitch levels in the form of a matrix, which is different from the one-dimensional energy superposition of the PCP feature, but instead uses a two-dimensional feature form - one of which represents the direction of the pitch level, and the other represents the direction of the octave. In this way, PCDM not only describes the energy distribution of each level from the perspective of the level, describes the characteristics of each frequency band (semitone) within each octave on the spectrum, but also effectively represents the chords and other important structural elements of the music, thus making PCDM, which is based on cognitive-distributed ideas, also has the characteristics of structural music features. Here the following algorithm is used to generate tone level distribution matrix features on the COT spectrum.

For a music signal at any time point , the pitch level distribution matrix characteristic $PM(t)$ is a matrix of size $t$ $M_m \times M_n$. The $M_m$ is the number of octaves, which, based on the experiments earlier in this paper, is set to 6. $M_n = 12$ is the number of levels, although this parameter can of course be set to some integer multiple of the number of levels that have practical music-theoretic significance. Given the real-valued spectral density $x_r[t,k]$ of the CQT spectrum transformed by $b = 120$ of the music signal at this moment, and the result $PO_t$ obtained by pitch-correction on top of this CQT spectrum, $PM(t)$ can be denoted as:

$$PM(t)[m,n] = x_r[t, 120m + 10n + PO_t] \tag{1}$$

where $m,n$ - satisfies $0 \le m \le M_m - 1$ and $0 \le n \le M_n - 1$.

In this way, $PM(t)$ gives a two-dimensional description of the overall spectral distribution of the harmonic structure of the superimposed voiced notes of the piece at moment $t$. In addition to capturing the overall sonic properties of the superimposed voiced notes at this time, $PM(t)$ also effectively characterizes the spectral distribution of the superimposed voiced notes within each octave. As a cognitively distributed music characterization, one of the more obvious weaknesses of $PM(t)$ in practical applications is its large data storage capacity. The $PM(t)$ defined above contains a time series of 6×12 data features at each moment, which is obviously a very large amount of computation when processing. In this regard, this paper proceeds to simplify it in two ways.

First, since $PM(t)$ is a continuous sequence of features in time, $PM(t)$ can be discretized according to the specific situation of the research task. Not all $PM(t)$ features at each moment are of equal importance, e.g., the $PM(t)$ of a segment connecting two notes is not only not required, but also interferes with the recognition of the elements. Therefore, in the actual use of $PM(t)$, this paper will generally do the detection of the starting point of the note first, and after obtaining the temporal information of each note, screen $PM(t)$ and select the $PM(t)$ at the time point where the spectral features are prominent and stable to be used.

Secondly, for $PM(t)$ at a certain time point, this paper will sparsify the data to remove some data points with insignificant spectral features, which can highlight the musical note characteristics described by $PM(t)$. This approach is as follows:

$$PM^*(t)[m,n] = \begin{cases} PM(t)[m,n] & \text{if } PM(t)[m,n] \text{ is top 3 within column} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where, $PM^*(t)$ - is the feature matrix after the sparsification process.

### *III. B. Methods of Ethnic Modal Analysis*

#### III. B. 1) Feature extraction

The tonal detection method based on PCDM distribution features proposed in this paper: for the target musical piece, this paper first performs a CQT transform at the b=140 level, and adopts a conversion frequency range of 28Hz~4194Hz, spanning an interval range of 7 octaves plus 1 minor third, i.e., A0~C8 on the tonal name, with a total of 88 semitones, which is just a range that covers all the standard piano keys that can be played. This range covers just about all of the range of musical notes that can be played on standard piano keys, and is truly the full keyboard pitch range. It should be noted that we did not use the 63 semitone A2~B7 melodic voicing shift range used in most of the experiments in this paper for the modulation detection task. The reason for doing so is that, as

modality is a global element, increasing the transformation range to full keyboard pitch can include more valid musical notes, thus better capturing the modal characteristics of the musical piece [20].

#### III. B. 2) Modulation identification

In this paper, a sequence of cognitively distributed features PM(t) is fed into a well-trained TANN network constituting a tonal recognizer for processing to obtain tonal recognition results. For the structure of the TANN network used is shown in Fig. 1. The structure of the TANN network can be changed in different situations, and specifically for the application of tonal recognition in this section, this paper uses a double hidden layer TANN network with a four-layer structure. Each input unit of the input layer is connected to all units of the first hidden layer, i.e., fully connected, while each unit of the first hidden layer is connected to three input layer units. The connection form between the first and second hidden layers is similar to that between the input layer and the first hidden layer, except that each second hidden layer unit is connected to five first hidden layer units; and each output layer unit is fully connected to the second hidden layer units. The design of such a connection form is based on the principle that higher level units should learn to make judgments based on lower level units in a wider range of time.

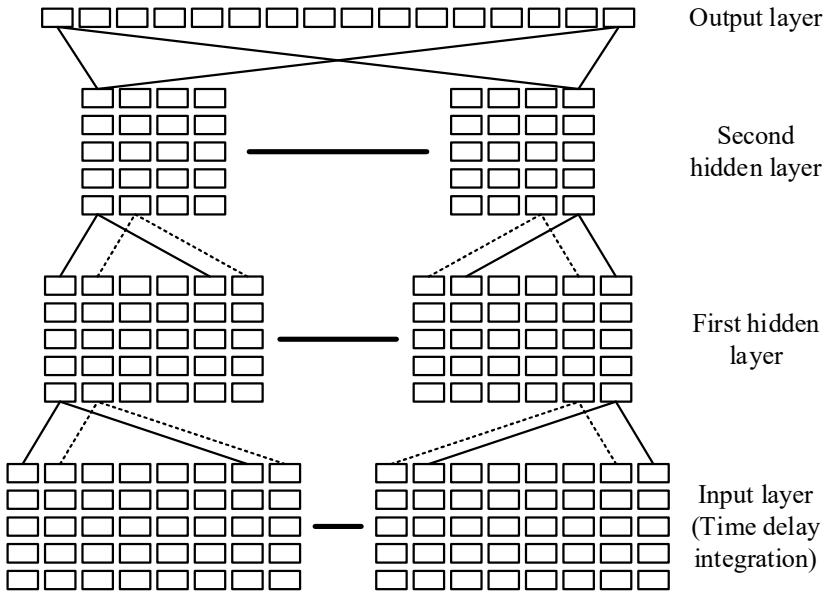


Figure 1: Mode Identifier TANN

#### III. B. 3) Modulation Smoothing

The method in this section is based on the detection of tuning based on the unit length of a tonal primitive, i.e., each section of a tonal primitive length is detected in its corresponding key, which inevitably results in the situation where neighboring sections are not detected in the same key. In the case of musical works in general, the change in key is far less drastic, and many entire pieces remain in the same key. Even though there are some works, especially modern pop music, in which the key is not static within a piece of music, and there are often transpositions or dissonances, the frequency of both transpositions and dissonances is very low, and it is rare to see multiple transpositions to different keys in the same piece of music. This requires post-processing of the detected tonal results - smoothing out detections with drastic local variations into the correct tonality, and ensuring that true points of transposition (off-key) are not ignored.

## IV. Characterization of the scale structure of folk tunes

### *IV. A. Experimental samples*

#### IV. A. 1) MIDI Chinese Piano Music Sample

The experimental samples are mainly monophonic melodies: Chinese folk songs in five-tone modulation and seven-tone modulation are interfered with to test the reliability of this paper's algorithm for the judgment of Chinese folk music modulation. Eight Chinese folk songs are selected as samples for the experiment, and the MIDI Chinese piano experimental samples are shown in Table 1.

Table 1: MIDI file format

| Serial number | Pentatonic | Heptonic |
|---|---|---|
| 1 | “Jasmine flower” | “Fighting in south mud bay” |
| 2 | “Liuyang river” | “Prairie love song” |
| 3 | “Stream” | “The decocour water” |
| 4 | “Chinese cabbage” | “I know you” |

**IV. A. 2) Data processing**

Since the names of Chinese folk tunes are composed of three parts, namely, principal tone name, style and type, with a total of 360 combinations, it is difficult to categorize them directly. In this paper, we adopt the way of categorizing the same house system to which the tunes belong, their principal tone names, styles and types to determine the final names of the tunes. For example, the D levy seven-sound elegant music modulation belongs to the G-gong system, with the main tone name of D, the style of levy, and the category of seven-sound elegant music. The categories are defined as shown in Table 2, for the first 3 of which the remaining 1 can be introduced from any two. After processing the database, the correspondence between the audio file of each track and its category number is obtained. For tonal identification, this paper uses two types of audio features, the chroma class feature and the constant Q transform (CQT) spectrogram. The basic principle of the Chroma feature is to map the spectral energy of the original audio signal into twelve octave-independent pitch classes, which can be represented by a twelve-dimensional vector, where each element corresponds to one of the twelve equal-tempered tone names, and contains the energy of that tone name over all the octave energies.

Table 2: definition of mode category

| Classification basis | Class name | Class number | Numbering |
|---|---|---|---|
| Homology system | C; #C(bD); D; #D(bE); E; F; #F (bG); G; #G (bA); A; #A (bB); B | 12 | 1-12 |
| Master sound | C; #C(bD); D; #D(bE); E; F; #F (bG); G; #G (bA); A; #A (bB); B | 12 | 1-12 |
| Style palace | Palace; Business; Angle A: feather | 5 | 1-5 |
| species | Five noises; Six words and a clear Angle; Six words and a palace; Seven songs; Seven noises; heptad | 6 | 1-6 |

CQT is a method to transform time domain information into frequency domain information, and its frequency resolution increases with increasing frequency, compared to Fourier transform, CQT is more suitable for decomposition of music signals. In this paper, we use the chroma_cqt function in Librosa library to obtain Chroma features, and use the cqt function to obtain the spectrogram, the frequency range in CQT is set to C1 to C8, and each octave is divided into 24 segments.

**IV. A. 3) Modulation templates**

The templates can be obtained in two ways: the first is based on the tonal definition and empirical settings, setting the tonic to 1, the out-of-tune to 0, and the partial to w. The second is to obtain the templates by counting the audio features in the database, e.g., when calculating the templates of the homophonic system, if the number of homophonic system of a piece of audio is s, then the normalized Chroma feature X will be shifted to left by a cycle of s bits and added to the template. For example, if the number of the homophone system of an audio is s, the normalized Chroma feature X will be left-shifted by s, then it will be added with the template, and the final template will be obtained by dividing the template value by the number of audio. The statistically based templates for modal categories are statistically derived in a similar way. After preliminary experiments, it is found that the template based on the modulation definition settings performs better than the template based on the statistical settings, so the template based on the modulation definition is used in the subsequent experiments. The specific values of the templates are shown in Table 3, and the highest correct rate of recognition is achieved when w is set to 1. When using the template of the same house system, the correlation coefficient between X and this template is calculated firstly, which can get the degree of matching between X and the C house system tuning, and after that the template is shifted to the right one place in the loop, and the correlation coefficient is calculated, which can get the degree of matching between X and the #C(bD) house system tuning. There are 12 correlation coefficients in total, where the largest result corresponds to the category to which the audio belongs to the same house system. When using the 5th, 6th and 7th tone modal type templates, the calculation is similar, and a total of 60 correlation coefficients are obtained, with the category corresponding to the largest value being the modal type. If the tonguing system template

is first used to obtain the tonguing system to which the audio belongs, and then the category recognition is performed according to the tonguing system and the category template, only six correlation coefficients need to be computed, but in the preliminary experiments, the correct rate of this method is lower than that of the recognition using the category template directly. In performing the principal tone recognition, a simple approach is taken by calculating the sum of the Chroma features of the last 500 frames of the audio to obtain a 12-dimensional vector, and the tone name corresponding to the maximum value is taken as the principal tone. Based on the recognized dominant t and the homophonic system s, the style of tuning can be inferred, with the Gong tuning when t is the same as s, the Merchant tuning when t is 2 semitones higher than s, the Horn tuning when it is 4 semitones higher, the Levitation tuning when it is 7 semitones higher, and the Feather tuning when it is 9 semitones higher.

Table 3: modulated template

| Formwork | Template value |
|---|---|
| System template | [1,0,1,0,1,0,0,1,0,0,1,0] |
| Based on the statistical template of the homopark system | [0.624,0.185,0.644,0.135,0.624,0.161,0.160, 0.721,0.164,0.669,0.098,0.285] |
| Pentatonic template | [1,0,1,0,1,0,0,1,0,0,0,1] |
| Six clear Angle template | [1,0,1,0,1,w,1,0,1,0,0,0] |
| The six sound and the palace template | [1,0,1,0,1,0,1,0,1,0,0,w] |
| Seven songs template | [1,0,1,0,1,0,w,0,1,0,1,w] |
| Seven sound relief templates | [1,0,1,0,0,w,1,0,1,0,0,w] |
| Seven-sound swallow template | [1,0,1,0,1,w,1,0,1,0,w,0] |

### *IV. B. Analysis of results*

#### IV. B. 1) Pentatonic modal analysis

The folk song “Jasmine Flower” in pentatonic mode was selected for testing. The calculated pitch and frequency of occurrence data were plotted. The pitch frequencies and scales of “Jasmine Flower” are shown in Fig. 2, and sorted according to the frequency of occurrence of each pitch, it can be seen more clearly that the main tones of the piece are all notes with a high frequency of occurrence. After preprocessing, it can be seen that the original pitch map was organized into 1 bar consisting of 5 pitches with a monotonic scale map.

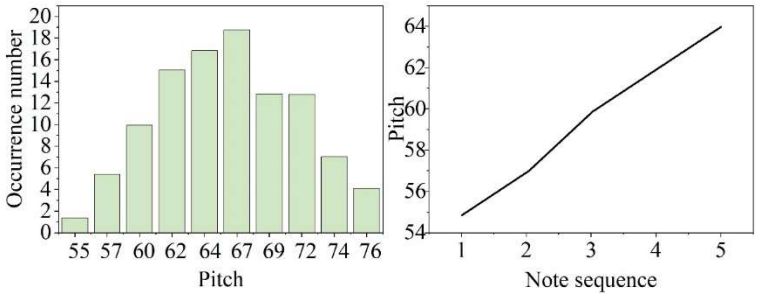


Figure 2: Pitch and scale

The calculations for Jasmine are shown in Table 4.

Table 4: The results of the jasmine flower

| Conditioning condition | Judging result | Modulated |
|---|---|---|
| Is it consistent with the national music mode template | It's in five tones | The gale is a five-tone modulation |
| Frequency highest sound (trunk) | 67 | |
| Wake | 67 | |
| The seventh level is related to the first level | No seventh level | |
| Characteristic tone group | Two plus four degrees | |

According to the above experimental results, we can conclude that through the statistics of the notes of "Jasmine", the number of occurrences of each note can be clearly obtained, and then the MIDI file is preprocessed to obtain a

pitch distribution map with obvious monotony, from which the scale of the melody of the work can be obtained, and finally the key is calculated as G pentatonic mode, which is consistent with manual judgment.

**IV. B. 2) Heptatonic modal analysis**

The pitch frequencies and scales of the piece are shown in Figure 3. Sorting each note according to its frequency of occurrence shows more clearly the number of times each tone appears in the work, with some tones appearing significantly less often than others. From the preprocessed figure, it can be seen that the original pitch map has been organized into 1 bar consisting of 7 pitches with monotonicity, and combined with the pitch frequency map, it can be found that the frequency counts of the 58th and 64th pitches are significantly less than the other pitches.

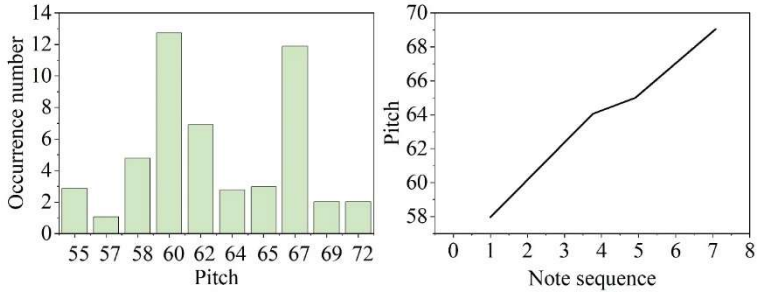


Figure 3: Pitch and scale

The calculation results of "Yes Our Confidant" are shown in Table 5. According to the above experimental results, we can conclude that through the statistics of the notes of "Is Our Confidant", we can clearly get the number of occurrences of each note, and then preprocess its MIDI file to obtain a pitch distribution map with obvious monotony, from which we can obtain the scale of the melody of the work, find out the partial tone, and finally calculate the key for Yanle G to seven tones, which is consistent with manual judgment.

Table 5: is the calculation result of "the person who is a person"

| Conditioning condition | Judging result | Modulated |
|---|---|---|
| Is it consistent with the national music mode template | The seven-tone template | The seventh tone of the song |
| Low frequency of frequency of frequency. | 58(bb jump), 64(f1) | |
| Partial gravity | 15.69% | |
| Frequency highest | 67(g1is) | |
| wake | 55(gme) | |
| The seventh level is related to the first level | Sophomore year | |
| Characteristic tone group | Two plus four degrees | |

**IV. B. 3) Results of sample testing**

The computational results of the experimental samples are shown in Table 6. From the experimental results, the recognition of Chinese folk music modes and non-Chinese folk music modes is basically correct, and the algorithm in this paper has high accuracy in recognizing Chinese folk music modes.

Table 6: The calculation results of the experimental samples

| Serial number | Experimental sample | Results |
|---|---|---|
| 1 | “Jasmine flower” | Five tone modulation (G) |
| 2 | “Liuyang river” | Five tone modulation (G) |
| 3 | “Stream” | Five tone modulation (C) |
| 4 | “Chinese cabbage” | Five tone modulation (F) |
| 5 | “Fighting in south mud bay” | The seventh tone of the song |
| 6 | “Prairie love song” | The lottery is the seventh tone |
| 7 | “The decocour water” | The seventh tone of the yan le |
| 8 | “I know you” | The seventh tone of the song |

## V. Conclusion

In this study, an automatic identification system of Chinese national tuning is successfully established through the combined application of tone level distribution matrix and temporal adaptive neural network. Experimental validation shows that the system achieves 100% accuracy in recognizing the tuning of eight samples of Chinese folk songs, which fully proves the effectiveness of the proposed method. In the specific analysis, “Jasmine Flower”, as a typical pentatonic modal work, has a significantly higher frequency of occurrence of the main tone 67th pitch than other notes, reflecting the scale structure characteristics of pentatonic modal. As for the analysis of the seven-tone modulation, the frequency of pitches 58 and 64 in “It's the One We Know” is significantly lower, and the off-key features are clearly recognizable, and it is finally accurately identified as Yan music G symbol seven-tone modulation.

The pitch level distribution matrix, as a two-dimensional feature description method, not only retains the advantages of traditional pitch level analysis, but also enhances the ability to characterize complex tonal structures through the introduction of octave dimension. The multilayer structure design of the temporal adaptive neural network enables the system to make tonal judgments over a wider time range, effectively avoiding the interference of local note changes on the overall tonal recognition. The template matching method based on tonal definition outperforms the statistical template in practical applications, indicating that the combination of theoretical foundation and algorithm design is important for improving the recognition accuracy.

This study provides a new technical path for the quantitative analysis of national tonal features in Chinese piano music composition, and plays a positive role in promoting the digital inheritance and innovative development of traditional music culture.

## References


[1] Lehmberg, L., BURTON, S., & REYNOLDS, A. (2018). Working with music from around the world. Engaging musical practices: A sourcebook for elementary general music, 169-198.

[2] Barton, G., & Barton, G. (2018). The relationship between music, culture, and society: meaning in music. Music Learning and Teaching in Culturally and Socially Diverse Contexts: Implications for Classroom Practice, 23-41.

[3] Waldron, J., Mantie, R., Partti, H., & Tobias, E. S. (2018). A brave new world: theory to practice in participatory culture and music learning and teaching. Music Education Research, 20(3), 289-304.

[4] Savage, P. E., Passmore, S., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. D. (2022). Sequence alignment of folk song melodies reveals cross-cultural regularities of musical evolution. Current Biology, 32(6), 1395-1402.

[5] Gustafson, R. I., & Gustafson, R. I. (2020). Folk Music and Dance: Imaginary Images of Modern Nationhood. Exploring Diasporic Perspectives in Music Education, 43-58.

[6] Richards, M. (2017). Tonal ambiguity in popular music’s axis progressions. Music Theory Online, 23(3).

[7] Gazor, M., & Shoghi, A. (2022). Tone colour in music and bifurcation control. Journal of Differential Equations, 326, 129-163.

[8] Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. Annals of the New York Academy of Sciences, 1423(1), 378-395.

[9] Li, Y. K., Loo, F. C., Osman, S., & Yao, P. Y. (2024). Exploring the multidimensional construction of cultural identity through music elements of Chinese Suzhou Pingtan. International Journal of Academic Research in Business and Social Sciences, 14(6).

[10] Bowling, D. L., Purves, D., & Gill, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. Proceedings of the National Academy of Sciences, 115(1), 216-221.

[11] Jiang, Y., & Zheng, M. (2024). The inductive effect of musical mode types on emotions and physiological responses: Chinese pentatonic scales versus Western major/minor modes. Frontiers in Psychology, 15, 1414014.

[12] Nan, N., & Guan, X. (2023). Common and distinct quantitative characteristics of Chinese and Western music in terms of modes, scales, degrees and melody variations. Journal of New Music Research, 52(2-3), 227-244.

[13] Liu, H., Jiang, K., Gamboa, H., Xue, T., & Schultz, T. (2022). Bell shape embodying zhongyong: The pitch histogram of traditional chinese anhemitonic pentatonic folk songs. Applied Sciences, 12(16), 8343.

[14] Xiyue, Z., & Rojanatakul, S. (2024). Piano Teaching of Chinese-Style Music Works. Journal of Roi Kaensarn Academi, 9(12), 3558-3565.

[15] Lu, N., & Dong, L. (2019). Research on the change of piano playing style from the national perspective. Frontiers in Art Research, 1(3).

[16] Zhang, Y. (2025). Eastern Rhythmic Foot and Western Colors: The Cross-Cultural Practice of Chinese Pentatonic Scales in Impressionist Music. Journal of Literature and Arts Research, 2(1), 116-126.

[17] Yu, M. C. D. (2014). Some Transformational Aspects of Pentatonicism in Post-tonal Chinese Music. International Journal of Contemporary Composition, 9.

[18] Yihe Jiang & Maoping Zheng. (2024). The inductive effect of musical mode types on emotions and physiological responses: Chinese pentatonic scales versus Western major/minor modes. Frontiers in Psychology,15,1414014-1414014.

[19] Zhu Yunli,Zhao Chengji,Zhang Yutong,Jian Meiru,Yang Haiwei & Ji Jiaming. (2024). Research on Emotional Conditioning Based on Chinese National Pentatonic Modes.

[20] Jiayin Sun,Haifeng Li & Lin Ma. (2011). A music key detection method based on pitch class distributiontheory. International Journal of Knowledge-based and Intelligent EngineeringSystems,15(3),165-175.