International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.7 (2016), pp.361-372
http://dx.doi.org/10.14257/ijmue.2016.11.7.36

Research on the Analyze Technology of Tuning Elements Based on the Theory of Cognitive Distribution

Jing Wang

Southwest university, Chongqing, China, 400715
616488304@qq.com

Abstract

The theory of cognitive distribution has been widely used in the field of artificial intelligence. This article takes music as example, and from the point of distributing view to bring out the hypothetical model of cognition that the researchers have on the music based the view of the characters of sound level distribution matrix (PCDM). Based on PCDM, combining with the technology of music signal processing, we propose an analyze technology framework of musical tuning elements based on cognitive distributed features. At last, this article carry out many comparative experiment of musical tuning testing upon the database constructed by many real music of different style, including the starting point detecting, the fundamental frequency estimating and the mechanisms of tuning smoothing.

Keywords*: Artificial Intelligence, Cognitive Distributed Theory, PCDM, Analyze Technology Framework, Tuning Testing*

1. Introduction

The research of thinking characters of cognitive distributed music originated from the engineering simulation of cognitive psychology on the theory of cognitive music. This thinking concentrate more on “distribution”, and most of the constituent elements of music can be distinguish recognized by analyzing the distribution pattern corresponding to the intrinsic properties of music [1]. However, this psychology statistics-based "distribution" and the distributed computing in computer science is not the same concept. The thinking of distributing cognition has been widely applied in many research fields already, such as the famous K-K algorithm. The performance of the thinking of cognitive distributed music is shown like below:

(1) It provides a new and comprehensive, accurate and high efficiency way to analyze the elements of musical tuning;

(2) In process of musical tuning elements analyzing, it has combined with the method of statistical pattern recognition of traditional vocals signal processing technology;

(3) During the process of analyzing and researching of the musical tuning elements, it also has played the advantage of experimental psychology-based approach, rules and model of music theory [2].

As the core part of musical theoretical concept system, musical tuning system has a close relationship with the most important elements of music. Sound level is a very important concept to decide the structure of musical tuning system and the relationship among its constituent elements [3]. Using music computing architecture to recognize and analyze the composition elements of the musical tuning architecture and the relationship among them, and this is an important task of automatic musical element analysis technology. This article based on the thinking of distributed cognition, proposes the PCDM cleverly through exampling the level and midrange octave forms of organization features of musical theory, and applied to analyze the global feature (musical tuning style) identification tonal system to distinguish the sound level of the organizational model.

ISSN: 1975-0080 IJMUE

2. Distributed Cognitive Characteristics and CDM

2.1. Distributed Cognitive Characteristics-Take the Music as Example

Distributed music based on cognitive characteristics of ideological thinks that, these learners(who have received the professional training before) stored the standard hearing template based on sound level distribution of a variety of musical elements in their auditory cognitive mechanisms, such as individual notes, complex chords, until all kinds of natural, harmony, melody tuning the sound level distribution template. When hearing the actual music, the learners will capture the sound level distribution of music segment they are hearing displayed by their sense of hearing, and compare it with the standard scale distribution template in the cognitive mechanisms [4]. Then find out the specific information of the corresponding elements according to the corresponding condition to complete understanding of the elements. This process is shown in the picture 1.

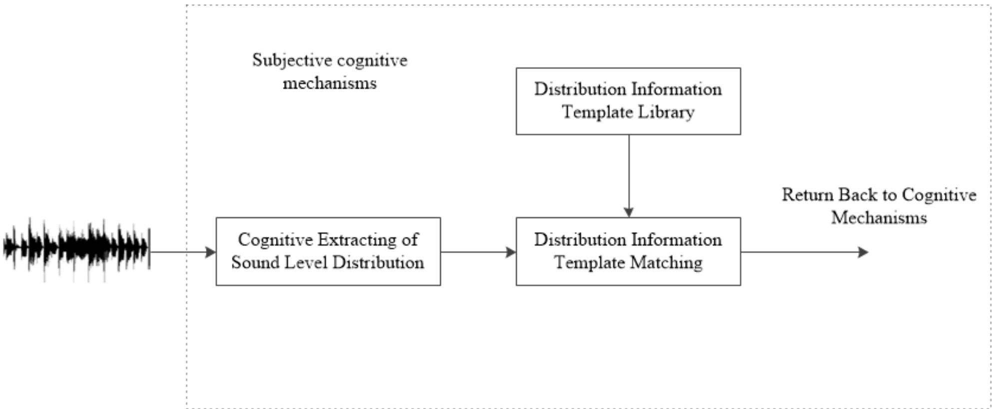

Figure 1. Distributed Cognition Theory of Music Foundation Thoughts

As the representative of the thinking based on cognitive distributed musical characteristics, because of the principle that the sound level contour capture audio characteristics is close to MFCC, and it is more suitable than MFCC to set music features' band, so it is widely applied in many works, and based on this, many modified sound level contour have been proposed [5]. All of these PCDM based on the sound level contour have the common feature, that is all of them is a one-dimension feature vector and take the sound stages or sound stages' integer multiple(such as the sound score in a octave) as the characteristic length. Each dimension of the feature vector stand for the energy superposition of all this has a roll-call of the scale within the specified range (usually no more than eight octaves). So the class of sound level contour can describe the distribution of the sound energy level. However, there still exist some problems in the sound level contour class features. Firstly, under the circumstance of describing the whole sound level distribution, sound level contour class features are lack of a description of the internal spectral characteristics of each octave, and only the internal spectral characteristics of each octave can reflect music's composition scale and structure directly. Secondly, sound level contour class features simply superimposed with a roll call of all the same energy scale makes some spectral characteristics to be "flattened", so the sound level contour class features only extract the music audio average spectral characteristics.

2.2. Pitch Class Distribution Matrix

Although sound level contour class features have been widely used in much music feature recognition problems, some of its shortcomings mentioned above will affect its effectiveness to count the features of music [6]. So this article proposes a Pitch Class Distribution Matrix (PCDM) based on the thinking of pitch class distribution, and is different from the sound level contour class features. PCDM describes the characteristics in the format of matrix, it is different from the sound level contour class features of one-dimensional energy superposition, and it use the feature form of two-dimensional, one dimension represents the direction of the sound level, another dimensional represents the octave's direction [7]. Like this, PCDM not only can describe the status of each sound level energy distribution from the perspective of sound level and describe each band in the spectrum of characteristic, but also can represent important structural elements of polyphonic music effectively. So that the cognitive thinking-based music characteristics like PCDM also has the features of structural music characteristics [8]. Here we use the algorithm like below to produce the features of PCDM on CQT spectrum.

To each musical signal at time point *t*, the features of its PCDM $PM(t)$ is a $M_m \times M_n$ matrix, M_m is the number of octave, according to the experiment, in this article, we set it as 6. $M_n = 12$, is the number of sound stages, of course, this parameter can also be set to the number of sound level music theory with practical significance of an integer multiple. Give the music signal at this time point, the density of CQT spectrum of real-valued spectral is $x_r[t, k]$, and it changes as the b is 120. Modify the pitch on this spectrum, and get the result of PO_t, and $PM(\mathrm{t})$ can be described like below:

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

Where $0 \le m \le M_m - 1$ and $0 \le n \le M_n - 1$.

Like this, $PM(\mathrm{t})$ gives out a two-dimensional description of the whole spectrum distribution of superimposed sound harmonic structure of music at the time point of *t*. In addition to capture the whole quality of superimposed sounding notes at this moment, $PM(\mathrm{t})$ also can describe the superposition of the spectral distribution of the sound note in the interior of each octave effectively. However, as the description of the musical characteristic, the based on cognitive distributed thinking, during the actually application of $PM(\mathrm{t})$, it also has the shortcoming of large data storage capacity. As it has defined above, $PM(\mathrm{t})$ is a time series with 6×12 data characteristics. Obviously, during the dealing process, the amount of calculation is very large. In this regard, this article simplifies it from two-pronged approach. Firstly, because the $PM(\mathrm{t})$ is a continuous feature sequence in time, we can discretize the $PM(\mathrm{t})$ according to the specific circumstances of the task. Not every $PM(\mathrm{t})$ at each time point has the same importance, such as the $PM(\mathrm{t})$ at the connection of two notes, they not only not required, but would interfere with the identification elements. So when we use the $PM(\mathrm{t})$ actually, firstly, in this article, generally I will do note onset detection. After obtaining time information of each note, I will filter the $PM(\mathrm{t})$, and choose the $PM(\mathrm{t})$ at the time point of spectral characteristics is prominent stable. Secondly, to the $PM(\mathrm{t})$ at some time point, this article will thinning the data, remove some data points, whose spectral characteristics is not obvious. This can highlight the musical features that the $PM(\mathrm{t})$ described. This practice is like below:

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

In the formula, $PM(\mathrm{t})^{*}$ is the matrix after the deal of thinning algorithm.

To the common musical works, this is a very rare case that with the same three roll-call, but belongs to different octaves of notes played at the same time. The possibility of three or more such notes played at the same time is much smaller. Therefore, this situation is not in this article. So, this paper uses this music theory rules to thinning $PM(\mathrm{t})$, sort the elements in the same column (belongs to the same sound level) of $PM(\mathrm{t})$, retain the three largest elements, and set other elements to 0. Although this algorithm is not optimal, and not excises the situation that the same three notes played at the same time, this has significantly reduced the size of the matrix already. Apart from these, when take the complexity of the thinning algorithm matrix into consideration, this thinning algorithm can be accept directly. At the same time of reducing the size of the matrix, thinning algorithm also by comparison of the 0 elements and non-zero elements to highlight the spectral characteristics superimposed notes described by $PM(\mathrm{t})$.

3. Musical Tuning Analyze Method Based on the Cognitive Distributing Features

3.1. Feature Extraction

Musical tuning is some high or low musical tone, circle around a central tone, which has a sense of stable (vocals), organizing them together by a certain principle, and form a system. From these we can see that: musical tuning is a system constructed by various elements of the musical tuning system, and it can express the features of the whole system, it is decided by the distribution of notes in the musical works. Musical tuning is the only one of the tonality property group reflects the specific structure's long property of the musical work not explicitly [9].

The testing method based on the distributing characteristics of PCDM proposed by this article is shown in the picture 2 like below:

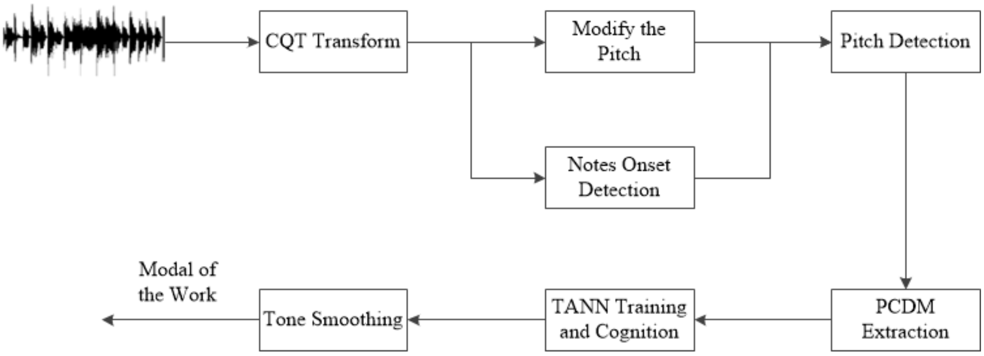

Figure 2. Musical Tuning Detection Process Based On PCDM

To the target music works, firstly, this paper carry out the CQT transform of *b=120*, and the range of transform frequency we adopt is $27.5Hz \sim 4186Hz$, across the pitch range of 7 octaves plus 1 small three degrees of sound range, that is the sound name of A0~A8, and there are 88 semitones. This range just covers all tone range of the standard piano keys that can be played, and is the full keyboard pitch range in the true sense [10].

Based on CQT transformation, this article uses the technology of pitch modification and notes onset detection to deal with the target musical spectrum. Shift the spectrum along the direction of frequency offset correction, and mark the location of the starting point of each note. After starting position of each known note, the method of this article is not for every frame subsequent processing, but after each successive eight for the starting point. This is because of that, from the point of view of the physical properties of sound note, after starting to sound decay time is the most significant period to express the spectral characteristics, so the method used by this article not only can reduce the actual amount of calculation, but also to capture the most effective spectral characteristics. This can form a discontinuous and CQT real value of the spectral density frame sequence after each note's start point, it is referred as the form of $x^{*}[t,k]$, t is not continuous in time.

After that, to the frame sequence of this discrete CQT spectrum, this article will estimate the pitch. To the data of each frame, this article will use the method of "*PreFEst*" to extract the fundamental frequency, and form the pitch sequence corresponding to CQT frame sequence. This is referred as the form of $F_0(t)$,where t is discontinuous in time. Meanwhile, we need use the fundamental frequency sequence to weighting CQT frame sequence $x^{*}[t,k]$: to each t belongs to $x^{*}[t,k]$, n_t is the nearest CQT spectral line to $F_0(t)$, where the t is not continuous in time. $y[t,k]$ is the spectrum weighted by $x^{*}[t,k]$, and the $G(n_t, \delta^2)$ is the Gaussian weighting function, and we can get $y[t,k]$ by the formula like below:

$$y[t,k] = \int_0^{\infty} G(\tau) \bullet x^{*}[t, k-\tau] d\tau \tag{3}$$

In the part of complex music, for the harmonic structure of every moment, the most important and largest contribution to the sense of hearing is the part of melodic, which is the root sound. The pitch of the root sound is measured by the fundamental frequency at that time. So this paper will test the fundamental frequency of time frame sequence at this time. Then weighting the CQT spectral according to the test results, this is equivalent to highlight the spectral characteristics of the root sound of the harmonic structure, in order to identify the tuning from the harmonic structure much more effectively [11].

Furthermore, extract the PCDM features from CQT spectral densities frame sequence $y[t,k]$, which is weighted by the PCDM producing algorithm, and it is referred as $PM(t)$.

3.2. Musical Tuning Recognition

In this paper, I transform the cognitive distributed signature sequence into the tuning recognizer for processing to get the result of the tuning recognizing, and the tuning recognizer is constructed by the well trained TANN net. To the TANN net structure that I used in this article is like below:

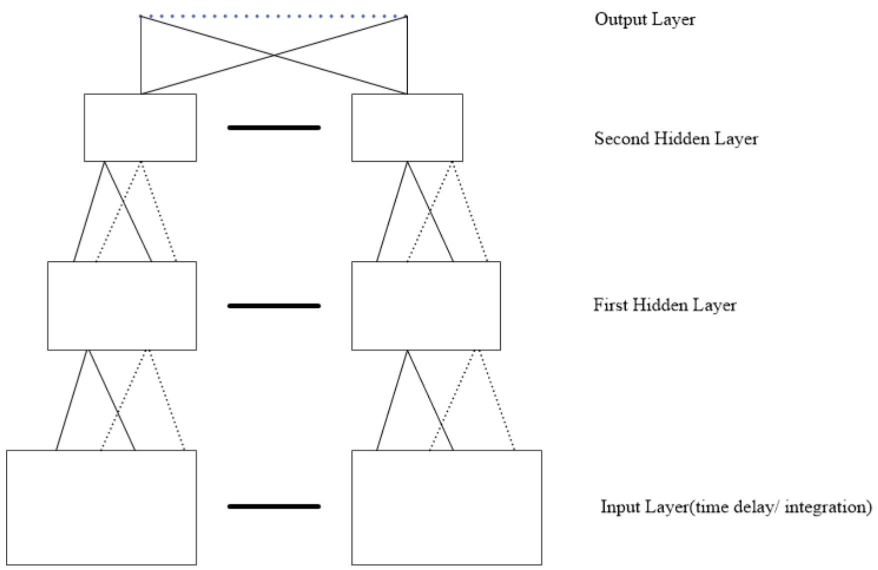

Figure 3.Tunning Recognizer TANN

To the net structure of TANN, it is changeable in different situation, to the application of tuning recognition in this section, I adopt this type of TANN, which is a four layer structure and has two hidden layers. Every input unit of the input layer connects to every unit of the first hidden layer, which is fully connected. And every unit of the first hidden layer connects to three units of the input layer. The connection type of first and second hidden layer is the same as the connection type of the input and first hidden layer, apart from that each of the unit of the second hidden layer connects to five units in first hidden layer. Each of the output unit is also fully connected. The design of this connecting type is based on the principle like below: the higher unit should learn to give out judgment based on the lower layer within a much wider time range. To these input characteristics of non-continuous time series, the TANN net process it with two steps.

Time delay- time delay mechanism includes two layers: within the harmony structure and among the harmony structure. Because of us only choose these eight data- after the onset of each note in this article. Therefore, we should do time delay accumulation of the length characteristics of the notes (harmonic structure). Meanwhile, because of the requirement of tuning recognition, we should also do delay among the notes. In order to demonstrate its necessity, we propose a new concept of tuning key unit. This article thinks that, to a musical work only with one tuning style, tuning key unit refers to the shortest length of music section that can determine its modal correctly. The length of the tuning key unit has something to do with the individual subjective judgment and objective environmental conditions. If take these person, who have received the professional training already, as the individual of subjective judgment, so the tuning key unit can be very short. If take these person, who can't understand the music, as the individual of subjective judgment, the tuning key unit may be infinite. The objective environmental conditions also can affect the tuning key unit, such as the noisy environment will increase the length of the tuning key unit. According to the classical musical theory, this article thinks that, whatever the cognitive objective is, the limited length of the tuning key unit should usually between three and four continuous normal chord, if shorter than this length, you may not have enough tuning elements (scales, chords) to identify specific tuning. So, during the process of tuning recognition, this article will import the concept of tuning element. Use the tuning element as the basic recognition unit of tuning to replace the traditional recognition unit, which takes the frame as the base unit: divide every three notes of the discontinuous characteristic sequence of *PM(t)* into a unit artificially, then each unit satisfies the condition of tuning elements (containing three consecutive harmonic structure). Like this, In order to enable the network to identify the tuning successfully, we should do time delay in every tuning elements- cumulate the characteristics sequence's time delay of the three harmonic structure. This is the time delay principle among the of harmony mechanism.

Integration-integration consists of two steps: time integrating and frame integrating. It should be noted that, no matter is time integrating or frame integrating, it's only for the harmony structure's internal continuous feature sequence, and no integration will be carried out among the three harmony structure.

The output of TANN net consists of 24 outputs; they present 12 major scales and 12 natural minors. If the output represents corresponding tuning is "1" or close to "1", and other output is close to "0", we think that the tuning has been detected. We adopt the training method of EEF to train the network.

3.3. Tuning Smoothing

The method adopt in this section takes the tuning elements as the unit length to detect the tuning, that is to say, corresponding tuning of each music section within the length of tuning element can be detected , it is inevitably that maybe the adjacent music section can detect out of different type of tuning. However, in general musical works, the change of the tuning is not so drastic like this; in the whole of many songs, they remained the same tuning. Although a part of the musical work, especially modern pop music, within the whole song, the tuning is not remains unchanged, it often turn or move away from the musical performance practices, no matter it is transpose or move away, its frequency is very low, and rarely appear the case that in the same piece of music repeatedly turned to the emergence of different tuning. This is what we will do after we detect the tuning. We should smooth this local test result with dramatic changes into the correct tuning, and make sure not to ignore the real turn (away) point.

Compared to the old tuning testing works, most of which usually only consider only a single tuning of musical works, the tuning testing algorithm has the advantages like below: we introduce the cognitive distributed features, which is close to the human's cognitive mechanisms. Use the TANN to simulate human's cognitive and identified mechanisms. And propose that by using the concept of tuning elements to test the target music's tuning section by section, and also take the tuning's change into consideration. Then give out the method of tuning smoothing under the circumstance of the tuning change is permitted, this can improve the efficiency of tuning recognition.

4. Experimental Results and Analysis

4.1. Choosing of the Experimental Sample

The data of tuning detecting experiment are these music clips that all come from different style and actually played, the audio frequency are all take the format of wav (22.05kHz sampling, 16 bit quantify and only has one sound channel). These data are all from five most common styles like the Table 1 has shown.

Table 1. Music Style

Style	Type	Length/section	Number of sections
Classical	instrumental music	90	90
Light Music	instrumental music	90	30
Jazz	Vocals and instrumental and part of them are on live	90	30
Pop/Rock	Vocals and instrumental	60	60
Chinese Folk	Vocals and instrumental	60	18

All of the callouts (tuning changing at every moment in all sections and the tuning change time point) referred above are all from standard music textbooks and tablature. Due to a small part of live version is improvisation, so it makes this section is different to the standard. In this regard, we invite professionals (students from the department of music) to mark the sense of hearing. More than two thirds of the music data is used to train the network, and the last one third is used to test.

4.2. Analyzing of the Result

To the training of the TANN recognizer, it adopts this two-stage training mode. Firstly, we run the professional notation software Overture 4.0 and combine with the mixing software HQ Orachestral VSTi to produce a tuning template of every major scale and natural minor. To each tuning we produce 20 such template, and we produce 480 this template in total, and use these template to train the network of TANN. Then, in the second step, put two thirds of the music sections (152sections in total) in this experiment into the network and complete to train the network. In the well trained TANN network, this article carry out the performance of testing of tuning detection.

Firstly, using the comparison experiment to test the performance of weighted in the PCDM using the fundamental frequency estimation, which is used to recognize the tuning, the result is shown in Table 2. Here, the main indicators to measure the fundamental mechanism for evaluating the performance are these: the accuracy of the test point of tone change in the music section (*pre*), the rate of recall (*rec*) and the accuracy of the music tuning detection. They can get through the formula below:

$$pre = (\mathrm{dm} - \mathrm{fpm}) / \mathrm{dm} \tag{4}$$

$$rec = (\mathrm{dm} - \mathrm{fpm}) / \mathrm{m} \tag{5}$$

$$acc = cdd / tmd \tag{6}$$

Table 2. The Comparison Experiment of whether it has Fundamental Frequency Estimation while Tuning Detection

	Our System	System Without the Fundamental Frequency Weighted
Total Time of the Music (tmd)/sec	6060.0	6060.0
Total Time of Tuning detected correctly (cdd)/second	4007.6	3854.3
Point Number of Tuning Change (m)	83	83
Tuning Change points detected (dm)	179	213
Tuning Change points detected incorrect (fpm)	98	144
Accuracy of Tuning Change points Detection (pre)(%)	45.25	32.39
Recall rate of the Tuning Change Points (rec)(%)	97.59	83.13
Accuracy of Tuning Detection (acc)(%)	66.13	63.60

Through this experiment we can see that, although the method of fundamental frequency detection has not reached an ideal correct recognizing rate, when it is used to strengthen (weighted) the spectral characteristics, it still have some improvements for tuning recognition method in this article- no matter the recognition of the tuning change points or the recognition of the whole tuning, it has an improvement in it. But this improvement is not very obvious, such as the accuracy of the detection of the tuning change point is still no more than 50%, and the accuracy of tuning detection only improve no more than 3%.

Then, to find out the function that the tuning smoothing mechanisms have on the whole method, we carry out the comparison experiment, and the result is shown in the Table 3.

Table 3. Whether it has the Smoothing Mechanisms

	Our System	System Without the Tuning Smoothing
Total Time of the Music (tmd)/sec	6060.0	6060.0
Total Time of Tuning detected correctly (cdd)/second	4007.6	3854.3
Point Number of Tuning Change (m)	83	83
Tuning Change points detected (dm)	179	213
Tuning Change points detected incorrect (fpm)	98	355
Accuracy of Tuning Change points Detection (pre)(%)	45.25	17.63
Recall rate of the Tuning Change Points (rec)(%)	97.59	91.57
Accuracy of Tuning Detection (acc)(%)	66.13	49.13

Through this experiment we can find out that, when process the continuous audio without using onset detection, there appears a lot of tuning changes point, the recognizing efficiency of the system become very poor, and the accuracy of the system tuning recognizing also reduce a lot. But to many system added error information at the start point manually, we can see that the error starting point information makes the accuracy of the tuning change point dramatically drop. Meanwhile, the rate of tuning recognition also decreased. So, from these we can see that the starting point of the system is very important.

Table 4. Result of the Comparison Experiment of Tuning Detection Refers to the Mechanism of Starting Point Detection

	Our System	System Without Starting Point Detection	System with Error
Total Time of the Music (tmd)/sec	6060.0	6060.0	6060.0
Total Time of Tuning detected correctly (cdd)/second	4007.6	2697.1	3445.7
Point Number of Tuning Change (m)	83	83	83
Tuning Change points detected (dm)	179	695	289
Tuning Change points detected incorrect (fpm)	98	629	214
Accuracy of Tuning Change points Detection (pre)(%)	45.25	9.50	25.95
Recall rate of the Tuning Change Points (rec)(%)	97.59	79.52	90.36
Accuracy of Tuning Detection (acc)(%)	66.13	44.51	58.86

Table 5. Experiment Detail of Tuning Detection Method

	Classical	Pop/Rock	Light Music	Jazz	Chinese Folk
Total Time of the Music (tmd)/sec	2700.0	1200.0	900.0	900.0	360.0
Total Time of Tuning detected correctly (cdd)/second	1824.3	717.4	631.6	597.8	236.5
Point Number of Tuning Change (m)	32	34	12	3	2
Tuning Change points detected (dm)	56	84	20	10	9
Tuning Change points detected incorrect (fpm)	24	51	8	8	7
Accuracy of Tuning Change points Detection (pre)(%)	57.14	39.29	60.00	20.00	22.22
Recall rate of the Tuning Change Points (rec)(%)	100.00	97.06	100.00	66.67	100.00
Accuracy of Tuning Detection (acc)(%)	67.57	59.78	70.18	66.42	65.69

At last, give out the experimental detail of the tuning detection method proposed in this section, it is shown in Table 5.From the result of classify we can see that, to this music section with vocals of other music style, the tuning detecting method performance much better on classical music and light music; these having vocals and rhythm intense, and with complex orchestration, the efficiency of tuning change point detection will become relatively poor. That is to say, the method we used in this article is has a certain degree of sensitivity to the music style, the main reason of this is that different styles of music express different characteristics in spectrum, this article is based on this method to process the music, and the spectrum characteristics of such classical music can be captured much more easier.

5. Conclusion

In this article, it propose an analyze technology of tuning elements based on the features of pitch distribution.

Firstly, from the point of distributed to bring out the hypothetical model of cognition that the researchers have on the music based the view of the characters of sound level distribution matrix (PCDM). PCDM not only can describe the status of each sound level energy distribution from the perspective of sound level and describe each band in the spectrum of characteristic, but also can represent important structural elements of polyphonic music effectively. So that the cognitive thinking-based music characteristics like PCDM also has the features of structural music characteristics.

Based on this, combining with the technology of music signal processing, we propose an analyze technology framework of musical tuning elements based on cognitive distributed features: use the technology of CQT, sound pitch modification, notes' onset detection, fundamental frequency estimation to do some front-end processing, then extract the features of PCDM, and transform the features sequence into to network of TANN to identify.

At last, based on the database consists of many different music styles to carry out the corresponding tuning detecting comparison experiment, such as onset detecting, fundamental estimating, tuning smoothing. According to these to prove the validity of the method based on the features of PCDM.

References

[1] A. Khallaayoun and Y. Huang, "Spatial Selective (S2) MUSIC performance analysis", Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation, Chicago, IL, **(2012)**, pp. 1-2.

[2] Z. I. Khan, M. M. Kamal, N. Hamzah, K. Othman and N. I. Khan, "Analysis of performance for multiple signal classification (MUSIC) in estimating direction of arrival", RF and Microwave Conference, 2008. RFM 2008. IEEE International, Kuala Lumpur, **(2008)**, pp. 524-529.

[3] T. B. Lavate, V. K. Kokate and A. M. Sapkal, "Performance Analysis of MUSIC and ESPRIT DOA Estimation Algorithms for Adaptive Array Smart Antenna in Mobile Communication", Computer and Network Technology (ICCNT), 2010 Second International Conference on, Bangkok, **(2010)**, pp. 308-311.

[4] Z. Rafii and B. Pardo, "REpeating Pattern Extraction Technique (REPET): A Simple Method for Music/Voice Separation", in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 1, **(2013)**, pp. 73-84.

[5] K. Lee and M. Slaney, "Acoustic Chord Transcription and Key Extraction from Audio Using Key-dependent HMMs Trained on Synthesized Audio", IEEE Trans. Acoust., Speech Language Process, vol. 16, no. 2, **(2008)**, pp. 291-301.

[6] M. Goto, "Frontiers of music technologies", Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015 IEEE Workshop on, New Paltz, NY, **(2015)**, pp. 6-6.

[7] S. Idowu, D. H. Hagos, W. B. Tesfay, A. Famurewa, J. Rana and K. Synnes, "NexTrend: Context-Aware Music-Relay Corridors Using NFC Tags", Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013 Seventh International Conference on, Taichung, **(2013)**, pp. 573-578.

[8] J. H. Su, C. Y. Wang, T. W. Chiu, J. J. C. Ying and V. S. Tseng, "Semantic content-based music retrieval using audio and fuzzy-music-sense features", Granular Computing (GrC), 2014 IEEE International Conference on, Noboribetsu, **(2014)**, pp. 259-264.

[9] X. Lin, D. Zeng, L. Cheng and J. Wang, "Study on the influence factors of music piracy in china based on SEM model", 2015 12th International Conference on Service Systems and Service Management (ICSSSM), Guangzhou, **(2015)**, pp. 1-6.

[10] L. Lu, D. Liu and H. J. Zhang, "Automatic mood detection and tracking of music audio signals", in IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 1, **(2006)**, pp. 5-18.

[11] X. Shao, N. C. Maddage, C. Xu and M. S. Kankanhalli, "Automatic music summarization based on music structure analysis", Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005, pp. ii/1169-ii/1172, vol. 2, **(2005)**.

Authors

Jing Wang, **S**he was born in May of 1978, and her nationality is hui. She has 16 years of work experience, in the four scientific research subjects. Music performance awards more than ten items. Now, she is studying for the religious music PhD degree in Southwest University. She also is a lecturer of the Music College in Yunnan Normal University, and her main research direction is philosophy and religious music.

