Relative Contribution of Frequency and Parameter Values to Selectivity for Interaural Correlation

Han Zhu[1], Hui Zhang[1], Song Wang[1*], Cong Zhang[1]

1 School of Mathematics & Computer Science, Wuhan Polytechnic University, Wuhan 430023, Hubei, China
Wangsongf117@163.com

Abstract. With the advent of the digital age, 3D audio is confronted with an important issues that how to compress multi-channel audio signal effectively. It is a significant way to explore the perceptual mechanism of human ear by researching the relative contribution of frequency and parameter values to selectively for Interaural Correlation (IC). Five values of IC that corresponding to five special sound width are selected in accordance with qualitative analysis of perceptual characteristics for IC: 1, 0.98, 0.94, 0.87 and 0.77. The audible frequency range of the human ear is divided into 24 bands based. A self-adaption test method contained the traditional test methods such as 1 up/2 down and 2AFC is taken, the test sequences are chosen as Gaussian white noise, and the experimental data is analyzed mathematically by the linear interpolation method. The experiment results demonstrate that: JND changes more significantly with frequency, and the minimum value occurs at about 700Hz and 1170Hz while the maximum values correspond to both side bands and the JND increases as the IC of the standard increasing. The experimental data and conclusions can provide theoretical support for the efficient compression of 3D audio signals.

Keywords: IC, JND, linear interpolation method, self-adaption test method, spatial audio coding

1 Introduction

Interaural Correlation (IC) also named correlative coefficient, it is an important parameter to describe sound stability and width in the human ear auditory system. The IC describe the degree of correlation between the left and right channels in stereo. The IC values range from 0 to 1, and the closer the IC value is to 0, the smaller the correlation between the left and right channels and the sound source is close to a surface source, it feels relatively divergent. The closer the IC value is to 1, the greater the correlation between the channels, the sound source is close to a point source and feels relatively concentrated.

The human ear has certain limitations to the perception of sound width. Only when the value of IC reaches or exceeds a certain threshold, the human ear will perceive the change of width of the sound. This threshold of perception is known as Just Notice Difference (JND). The smaller is the JND value, the stronger is the sensitivity of the ear, it is the easier to perceive the change of width of the sound. In order to explore the sensitivity of human ear to spatial parameter localization, namely the ability of human ears to distinguish IC, some scholars have studied the JND of IC by various experiments. The frequency of sound signal, IC parameter value and experimental environment will affect the value of IC. The scholars have carried out various measurements and analyses on these influencing factors.

The theory of double ear masking has proved the importance of IC in the ears listening [1], but there are only a few articles to study the human ear's perception of the IC, and most of them have only made the sensitivity experiment of the IC reference value at 1, some researchers have experimented with IC reference values of 0 or -1 occasionally, the three reference value experiments have only begun in recent years. It is very difficult to analyze advantages and disadvantages of experimental methods used by these different binaural masking models. Because their inner forms often change with the change of incentive

---

* Corresponding Author

parameters.

As early as 1959, Pollack and Trittipoe studied the effects of the parameter values of the sound length, frequency, sound source level and binaural correlation IC on JND values [2]. They used a Gaussian noise as the test sequence with length of 10~1000ms, frequency range of 100~6800Hz, sound pressure level of 50~ 90dB, and IC parameter range of 1~0. Finally, they found that when the reference value of IC is 1, the IC value of JND is about 0.04, when the parameter value of IC decreases to 0, the JND value of IC increases to about 0.44.

In 1981, Gabriel and Colburn also used Gaussian noise as sound source to test the influence of the bandwidth of the sound source and sound pressure level on the JND value of the binaural correlation IC [3]. The center frequency of the sound source is 500Hz, the bandwidth range is 3Hz~4.5 kHz, and the spectrum intensity of sound source is 75dB and 39dB. The results show that IC's JND value is maintained at a parameter value close to 0.004 when the reference value of IC is 1 and the bandwidth is less than 115 Hz. When the bandwidth is greater than 115Hz, the JND value of IC is monotonically increasing to about 0.04. When the reference value of IC is 0 and the bandwidth is changed from 3 to 115 Hz, the JND value of IC changes from 0.7 to 0.35. When the bandwidth is greater than 115 Hz, the IC's JND value is a constant that is close to 0.35. When the reference value of IC is 1, the JND value of IC decreases with the increase of spectrum intensity, and when IC reference value is 0, the JND value does not change.

In 1993, Cox and others used music as a source signal [4]. They simulated the loudspeakers at the concert hall and use sound processing software to process analog sounds directly into early reflection and reverberation combinations of anechoic recordings as test sounds. Finally, they found that when the reference value of IC was 0.33, the value of JND was 0.075±0.008.

In 2002, Okano also used the silencing music recording to test the JND value of IC [5]. In particular, the test sound is expressed in the form of 1-IACCE3 (E represents the direct sound of 0~80ms, and 3 represents the triple band in the center of 50, 1000 and 2000Hz). Direct sound, early reflection and reverberation are generated through multiple loudspeaker simulation, where the level of early reflection is constantly changing. The experimental results show that when the reference value of IC changes from 0 to about 0.8, the value of JND is 0.065±0.015.

In 2009, Lüddemann generated all of the test sequences in matlab, and each test sequence was stitched together by a pair of double channel Gaussian band noise segments (bandwidth of 100-2000Hz) of different IC values [6]. The IC value of each fragment in the test sequence is set by mixing two orthogonal noise sources in a certain proportion. In order to calculate the validity, only the right channel is processed in this experiment. The standard deviation of the maximum IC value can be accepted by the experiment: standard deviation is 0.003 when ρ is equal to 1 and -1; standard deviation is 0.03 when ρ is equal to 0. IC value at the threshold between them is obtained by linear interpolation method, only when the deviation between test sequence and target sequence is under the expectation values, the sequence can be used, and otherwise the sequence will be discarded and replaced with a new sequence. It was found in the experiment that when IC reference value is 1, JND is 0.034; when IC reference value is -1, JND is 0.105; when IC reference value is 0, the value of JND at positive value is 0.547, while the value of JND at negative value is 0.667.

In 2013, according to the frequency band characteristics of human ear perception, Cao Cheng divided the audio signal into 24 subbands by the Bark frequency division method, the value of IC is determined by the deviation angle of sound source boundary and human ear, five test sound take the corresponding IC value of deviation angle of 10º, 20º, 30º, 40º, 50º respectively [7]. The testers subjectively distinguish the sound source width difference between the test sound and the standard sound, by which the sensitivity of human ear to binaural correlation IC is determined. The results of this experiment show that the IC parameter's JND sensing curve contains three peaks, which are located in low frequency, intermediate frequency and high frequency. The area of the JND value is relatively small in the three peaks, and it is easy to detect the variation of the width of the sound field in these bands.

Statistics of the JND test results of IC in several classical literatures show that most researchers study for the narrow band, only Dajani and Picton made a low and medium frequency band IC sensitivity test for 0~8000 Hz. Judging from the JND test of IC, only JND of three reference IC values were tested by Boehnke in 2002 and Luddemann in 2009. Even with the same IC reference value, the value of JND obtained by scholars from different ages is very different. Therefore, it is necessary to make a complete experiment for all bands and different IC reference values, to establish a perceptual model of JND and

frequency of IC and IC, to explore the most basic physical phenomena of human ear.

IC is an important parameter to describe sound stability and width in ear auditory system, and an important parameter of spatial sound source localization, the JND of IC reflects the sensitivity of human ear to the change of sound width. Most researchers study for narrow band, and different scholars get the different JND by the same IC reference value, so the research contents of this paper mainly include:

(1) Improve the current listening system to obtain an audio listening system for IC perception

In this paper, an audio listening software based on windows system is improved, aiming at the demand of the current dual-ear correlation IC perception test experiment.

(2) The JND data of the full band is measured with different IC parameters

The improved listening system can measure the JND value of IC in the full band range under different IC parameters, and the JND data obtained is more comprehensive and complete than before.

(3) Mathematical analysis and modeling of IC's perceptual characteristic JND

Using mathematical fitting method, the experimental data measured in (2) are analyzed, fitted out about surface of the JND-frequency-sound image width of IC, so the relationship between JND value of IC, frequency and IC parameters can be analyzed more specifically to provide theoretical support for multiple channel audio signal efficient compression.

2 Experimental Method

2.1 Subject

All listeners of this experiment have passed the routine hearing examination of the hospital. Prior to this test, everyone have received professional listening training and many subjective listening tests. Prior to this experiment, all the listeners should be trained professionally to understand the test principles of this experiment. In order to obtain more accurate experimental data in the subjective experiment, the listeners should be screened twice to ensure the accuracy of the experimental process. Through the screening of the listeners, seven subjects were selected including four men and three women, aged from 21 to 27. They all are students at school.

2.2 Test Environment

In order to reduce the negative effect brought by environmental and equipment to the tester and get more accurate experimental data, the test is done in professional audio lab experiment. The subjects wear the same professional headphones, and complete the experiment on the same computer. The hardware and software used in the experiment are shown in Table 1.

Table 1. Test environment / equipment

Equipment/Environment	Parameter
CPU	Intel Core i5-4590, 3.30GHz
RAM	4GB DDR3 1600MHz
OS	Windows 7 (64bit)
Externa Soundcard	Creative Sound Blaster X-Fi HD
Headphone	Sennheiser HD380Pro
Sound Level Meter	Smart AR824
Software	Visual Studio 2005, Adobe Audition, Matlab 2014

The listening device used in this experiment is headphone, which can reduce the influence of surrounding environment, so as to ensure the efficient experiment process and reliable experimental results.

2.3 Test Signal

According to the spatial psychological acoustics, there are many factors influence the JND value of the two ear correlation IC, such as signal type, frequency of sound signal and sound image width.

Signal type. Gaussian white noise, whose sampling frequency is 96 kHz, length is 300ms, and the sound pressure level of the audio signal is controlled by the sound level meter, ensuring that the sound output of the audio signal in headphone is 75dB.

Selection of IC reference values for test signal. Set S1, S2 is the original left and right sound source, L, R is the left and right sound channel after mixing, and then the synthesis formula of L, R:

$$L = S_1 \cos\alpha + S_2 \sin\alpha \tag{1}$$

$$R = S_1 \cos\beta + S_2 \sin\beta \tag{2}$$

In order to meet the symmetry of the right and left mixing, set $\alpha + \beta = \dfrac{\pi}{2}$, considering 'unrelated' and energy equal, we can get:

$$IC = \frac{\text{cov}(L,R)}{\sqrt{\text{cov}(L,L)} \bullet \sqrt{\text{cov}(R,R)}} = \cos(\alpha-\beta) = \cos(\frac{\pi}{2}-2\alpha) = \sin(2\alpha) \tag{3}$$

Each IC value is corresponding to an angle value 2α. It is assumed that the corresponding angle between the sound source boundary and horizontal direction is called deviation angle. The vertical direction is defined as 90º, horizontal direction is defined as 0º. As shown in Fig. 1.
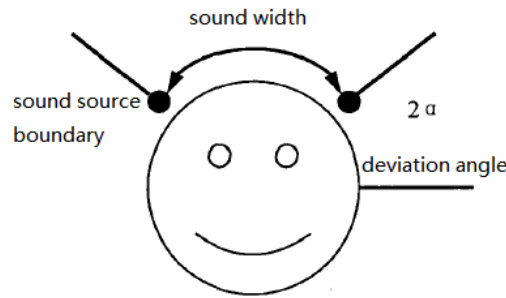

Fig. 1. Schematic diagram of deviation angle

In the actual test, the width information of the sound is changed to the boundary information of the sound, which is easier to judge in the test. Deviation angle are 90º, 90º, 70º, 60º, 50º respectively, the corresponding IC value is 1, 0.98, 0.94, 0.87, 0.77, which are five reference values in the test.

Band allocation. In addition to exploring the relationship between IC values and JND, the relationship between frequency and JND is also need to be explored. According to the frequency band characteristics of human ear perception, the Bark frequency band division method is used to divide the audio signal into 24 bands, and the frequency coverage is 20Hz~15500Hz [8]. Gaussian white noise is deal with band pass filter, the narrow-band Gaussian white noise with 24 bands are obtained: 0~100Hz, 100~200Hz, 200~300Hz, 300~400Hz, 400~510Hz, 510~630Hz, 630~770Hz, 770~920Hz, 920~1080Hz, 1080~1270Hz, 1270~ 1480Hz, 1480~1720Hz, 1720~2000Hz, 2000~2320Hz, 2320~2700Hz, 2700~3150Hz, 3150~3700Hz, 3700~4400Hz, 4400~5300Hz, 5300~6400Hz, 6400~7700Hz, 7700~9500Hz, 9500~12000Hz, 12000~ 15500Hz.

Generation of test sequence. Audio signal is generated by the Audition, basing on Bark band division method. Audio signal is Gaussian white noises with 24 specific center frequencies, their sampling frequency is 96 kHz, their length is 300ms, their sound pressure level is 75dB, controlled by a sound level meter. Each test sequence is randomly combined of the reference audio and test audio. IC values of the reference audio and test audio are different, their length is 300ms, and there is a 250ms mute interval between them, so each test sequence of a total length of 850ms. Both sequences are stereo signals, which replicate the filtered single channel signal into a double channel. The stereo test sequence of different binaural correlations is obtained by changing the acoustic image width.

Configuration of experimental parameters. The listening system used before in this laboratory is designed based on the perception characteristics of ILD and ITD. Therefore, before the experiment, we would like to improve audio listening software based on Windows system, and preset the parameters in the configuration file.

```
    [Mode]
    refMode=3
    testMode=3
    strFrequence=50,150,250,350,450,570,700,840,1000,1170,1370,1600,1850,
2150,2500,2900,3400,4000,4800,5800,  7000,8500,10500,13500
    strStep=0.9,0.9,0.03,0.02
    nDuration=300
    nFs=96000
    fVariable=0.1
    fReference=1.0
end.
```

The meaning of each parameter and the value assigned to it is:

· When refMode and testMode are set to 3, the system calls SequenceICCreat.

· FReference = 1.0 indicates that after the system calls SequenceICCreat, the reference sequence will be generated with an IC value of 1.0.

· FVariable = 0.1 indicates that the test sequence will be generated after the system calls SequenceICCreat, and its IC value is 0.9.

· StrFrequence represents the frequency of the test sequence in this group, and the system will select Gaussian white noise from the source folder according to this frequency to generate the test sequence.

· StrStep is parameter of step change, and strStep directly influences the change of IC value in the test sequence.

After configuring the parameters required for the experiment, the listener can make experiment with the system.

2.4 Methodology

Experimental sample. First, the listener listened to two sequences of fixed binaural relativity (The IC value of the reference sequence is 1, the reference value of the test sequence is 0.8), and then they determine the sound image range of which sequence is too large. The excel table will record the results of each judgment. If one can reach 80 percent of the correct rate (with the latest 20 results) in the number of times (50 times), he or she can filter. This round of screening tests subjects' ability to perceive changes in the sound image caused by changes in IC, and also tests the patient's patience. All subjects will be selected after the screening.

This experiment contains several tests of different IC values, and each group also needs to be listened on multiple sound frequencies. To avoid causing test fatigue, it takes a break after several frequency points to be tested. It takes about two hours to complete a test.

Experiment method. In the actual test process, it is easy for listener to distinguish the different sound image width at the beginning of the test. But as the IC difference between the two audio signals becomes smaller and smaller, it becomes more and more difficult to judge the results. As the test amount of this experiment is relatively large, in order to simplify the experimental process, improve the test efficiency, the psychological adaptive testing methods of 1 up/2 down [9] and two alternative forced-choice (2AFC) are adopted [11].

The 1 up / 2 down is one of the Transformed up-down psychological adaptive test methods. Since each group needs to test many times, test results of each round will directly affect the test sequences that are generated in the next round. The process of 1 up/2 down psychological adaptive test method is shown in Fig. 2.
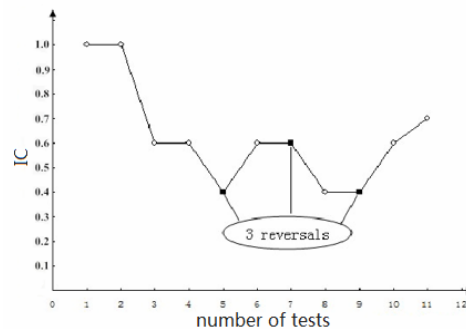

Fig. 2. 1 up/2 down test process diagram

Each point in the diagram represents a test result, the hollow point indicates that the judgment is correct, and the solid point indicates that the judgment is error. The IC value will decrease when two consecutive judgments are correct. The IC value increases as soon as an error judgment occurs. A total of 3 inversion points appeared in the 12 rounds of listening test of Fig. 3. Inversion refers to the process of change of IC value from increase to decrease or decrease to increase, one change is the equivalent of a reversal. Before the experiment, we need to set the threshold for the inversion point to set when to finish the test.

The system reduces the IC value of the current test sequence to a step length when the tester correctly evaluate twice. It makes the sound image width of the test sound and reference sound in the next listening sequence to be more similar. As long as there is a round of error in judgment, the system will increase the IC value of the test sound based on the current step size. In this way, during the multi-round testing, the IC value of the test sound is constantly adjusted according to the testers' judgment, which gradually approximates the value of the JND which is to be measured. 1 up / 2 down psychological adaptive test method is based on the preset inversion point threshold to determine whether to end the test of the group. We set up the initial step length of JND based on the experimental results we have obtained and the experimental data obtained from the small range. When the IC reference value is 1, 0.98 and 0.94, the initial step size of JND is 0.3, and when the IC value is 0.87 and 0.77, the initial step size of JND is 0.2.

2AFC experimental method requires the subject to hear the test sequence in each round, must be based on their own subjective feelings, within the specified time to choose their own voice range is too large to say the first half or the second half. Regardless of the final outcome of this judgment is correct or not, must make a choice. If you do not make a selection, the system will regenerate a similar test sequence based on the current parameter value to let the listener re-judge.

The 2AFC experiment method required that listeners listen the test sequence in each round, and choose the sound they think the rang of sound image is large from the first half or the second half based on their own subjective feelings within the allotted time. Whether the final result of this judgment is correct or not, a choice must be made. If no selection is made, the system regenerates a similar test sequence based on the current parameter values to make the auditors re-evaluate.

Suppose that the IC value of the test sound is IC_{test}, the IC value of the reference sound is IC_{ref}, and the change step is IC_d, then we can get:

$$IC_{ref} = IC_{test} + IC_d \tag{4}$$

Step 1: Set the value of IC in the configuration file sett.ini (set IC is 1, 0.98, 0.94, 0.87, 0.77 respectively) to set a fixed sound image width for the reference sound. For example, in the first set of experiments, the IC value of the reference sound is set to 1 ($IC_{ref} = 1$), indicating that the reference sound is a point source in front of the listener. If the initial IC value of the test sound is set to 0.9 ($IC_{test} = 0.9$), the sound width of the test sound is divergent compared to the reference sound.

Step 2: After the configuration is completed, we enter the listening system and run. First, the hearing normal of the tester is determined through the listening training phase, then the formal listening stage is followed. The tester selects the frequency X of the audio signal to be tested in the drop-down box at the back of the test sequence. After click ok, the system will generate the reference and test sound, according to the selected test frequency X, IC value of reference sound (fReference) in the configuration file in step 1 and IC value of test sound (fTest = fReference - fVariable), and randomly combine them into a test sequence. After that, the generation of each test sequence is determined according to the changed step strStep.

Step 3: Test the listening sequence generated by step 2, then select the first half and the second half with bigger sound image range in the specified time. Set three global variables Nr, Nw and Reversals with initial values zero, where Nr and Nw respectively indicate the correct number and wrong number of results, and Reversals indicates the number of inversion. If the tester judges correctly first time of this round, then set Nr=1, Nw=0. If the second judgment is correct, then Nr+1, that is Nr=2, then according to the principle of 1 up / 2 down psychological adaptive test method, reduce the value of IC_d, and reset Nr and Nw to 0, and then determine whether inversion happens. If inversion happens, the value of Reversals is incremented by 1 and the value of current Reversals and IC_d are saved, and goes to step 4. If inversion does not happen, then return to step 2, generate a new test sequence according to the current IC_d, then go to the next round of testing. If judgment is wrong, set Nw=1, Nr=0, the value of IC_d is increased, and Nr and Nw are set to 0 at the same time, then determine whether inversion happens. If inversion happens,

the value of Reversals is incremented by 1 and the value of current Reversals and IC_d are saved, and goes to step 4. If inversion does not happen, then return to step 2, generate a new test sequence according to the current IC_d, then go to the next round of testing.

Step 4: The inversions times Reversals is preset to a value of L and determine whether the current reversals times has reached L times. If the reversals times has reached L times, we should go to step 5, otherwise we should go to step 2.

Step 5: We calculate the average of the last n times IC_d values of the inversions times, where n is the preset value, and then we can get the value of the JND in the X band.

After the first frequency range test has been completed, we choose the remaining 23 frequencies in the drop-down option box to test. After the first group tests is completed, we open again the Setting.ini file and modify the values of fReference and fVariable, in the same way to complete the following four groups of experiments (fReference is equal to 0.98, 0.94, 0.87, 0.77). Each time, the test results and test data are stored in the corresponding excel table.

In step 3, the reduction and increase of IC_d are described by the following formula:

$$IC_d = IC_d * index - linear \tag{5}$$

$$IC_d = IC_d / index + linear \tag{6}$$

Where index is the exponential change parameter, linear is the linear change parameter, their value is obtained by the strStep in the configuration file. The reduction and increase of the step size IC_d are determined by Equation 5 and Equation 6.

According to the above experimental method to develop listening test software based on Windows system, the first things is to set the configuration file parameters of this listening test software, and then tester enter the training interface to judge correctly, finally tester enter the formal listening stage. Tester choose the value from the drop-down box for listening test, and then complete listening test of each frequency under current IC reference respectively. After this group test, we modify the reference value of IC and the initial step of JND in the configuration file for the next set of tests. All test results are saved in the excel table in order to be dealt with and analyzed in the later period.

3 Results and Discussion

After four months of testing, seven subjects completed the binaural correlation IC-sensing characteristics test under five reference values and 24 bands, and finally we calculated the average value of the seven subjects stored in the excel sheet.

3.1 The Relationship between JND and Frequency

In order to analyze the relationship between JND and frequency, we use a line chart showing the relationship between the JND value and frequency as shown in Fig. 3.
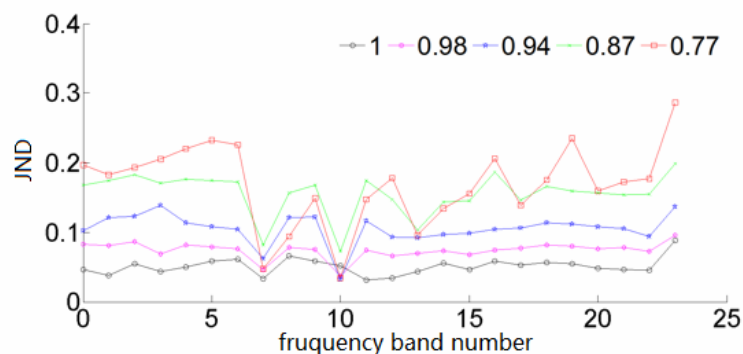

Fig. 3. The JND value of IC varies with frequency

From Fig. 3, we can see that:

(1) On the whole, the JND of the same IC in different bands is different, that indicates the impact of IC

value on the spatial sound source location in the various bands are not the same, even in some frequency bands are very different, and so we need to analyze it with different frequencies.

(2) The JND curve of the IC parameter contains several peaks, and most of the peaks are located in the middle or high frequency. The JND value between the peaks is relatively small, which indicates that the human ear is relatively easy to detect the variation of sound and image width in these bands

(3) Locally, the JND value is minimal when the frequency is 840Hz and 1370Hz. It indicates that it is the most sensitive for human ear to detect the variation of sound and image width at about 840Hz and 1370Hz. About 1000HZ frequency is a special frequency. The minimum value appears at 1000 Hz in the JND test of ITD (Interaural Time Difference). The maximum value appears at 1000 Hz in the JND test of ILD (Interaural Level Difference) [12].

(4) With the decrease of IC value in reference, the JND value increases and the whole line moves upward. It is shown that when the IC value is larger, the human ear can easily perceive the change of the sound image width, and as the sound image width increases, it becomes more and more difficult for the human ear to perceive the changes in the sound image width. When the IC value is close to 1, the change of JND is more stable, and when the IC value is reduced to 0.87 and 0.77, the fluctuation of JND value is obvious. It is shown that when the IC value is large (the sound source is close to the point source), the perceived ability of the human ear to the sound image width is relatively stable; when the IC value is small (the sound image width is large), the human ear has an unstable perception of the sound image width.

3.2 The Relationship between JND and IC

In order to analyze the variation law of JND and IC, the six bands of 1, 5, 9, 13, 17 and 21 were selected and analyzed. The results are shown in Fig. 4.
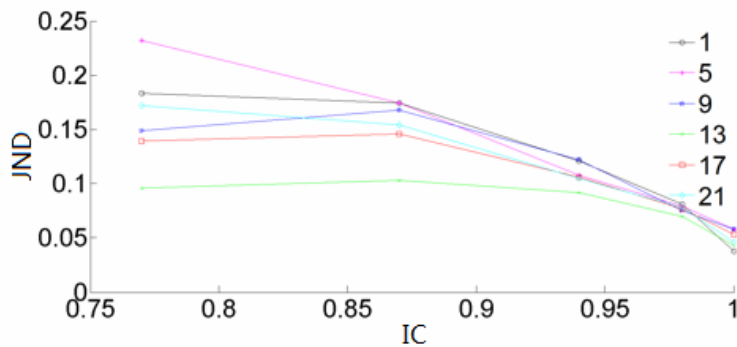

Fig. 4. The JND value of IC varies with interaural correlation

From Fig. 4, we can see that:

(1) On the whole, the JND value decreases with the increase of IC value in reference. It is shown that when the IC value is larger, the human ear can easily perceive the change of the sound image width. As the IC value decreases, the sound image width increases, and the human ear becomes more and more difficult to perceive the change of sound image width.

(2) Locally, when the IC value is close to 1, the difference of JND value in different frequency bands is very small. When the IC value reduces to 0.87 and 0.77, the fluctuation of JND value is obvious. It is shown that when the IC value is large (the sound source is close to the point source), the perceived ability of the human ear to the sound image width is relatively stable; when the IC value is small (the sound image width is large), the human ear has an unstable perception of the sound image width.

3.3 The Relationship between JND, Frequency and IC Value

To study the relationship between JND of IC, frequency and IC value, for more frequency and IC reference value, we interpolate the JND of IC, and get a three-dimensional surface map of the JND-Frequency-IC value with more precise interaural correlation IC. At present, JND data has been measured, and the most critical step is to select the appropriate interpolation points for the sound image width reference IC and frequency in the experiment.

(1) The selection of interpolation point of the IC reference value

The IC values selected for the experiments are 1, 0.98, 0.94, 0.87 and 0.77. According to the JND curve of interaural correlation perceptive characteristic, with the decrease of IC value, the sound image width increases, and JND increases accordingly. When the IC value is larger (the sound width is smaller), there are more interpolation points. When IC value is small (the sound width is smaller is larger), there are fewer interpolation points. Therefore, the interpolation points of the sound image widths are 0.99, 0.96 and 0.91 respectively.

(2) The selection of the interpolation points of frequency

24 bands is selected by the test signal in the experiment, which corresponds to 24 Bark values. So the result of the test is the perceptual characteristic of center frequencies of 24 Bark frequency band. The JND perceptual characteristic of the interaural correlation IC at the frequency point between two Bark bands is similar to the JND perceptive characteristic of the two Bark bands themselves. To make the final surface smooth enough, the interpolation points we selected are the boundary frequency points of each Bark band. The selected frequency interpolation points are 100Hz, 200Hz, 300Hz, 400Hz, 510Hz, 635Hz, 770Hz, 920Hz, 1085Hz, 1270Hz, 1485Hz, 1725Hz, 2000Hz, 2325Hz, 2700Hz, 3150Hz, 3700Hz, 4400Hz, 5300Hz, 6400Hz, 7750Hz, 9500Hz and 12000Hz.

The selected frequency interpolation points are the boundary frequency points of each Bark band. The interpolation points of IC values are 0.99, 0.96 and 0.91 respectively. In this paper, the linear interpolation method is adopted to fit the 3D JND surface of IC parameters, as shown in Fig. 5.
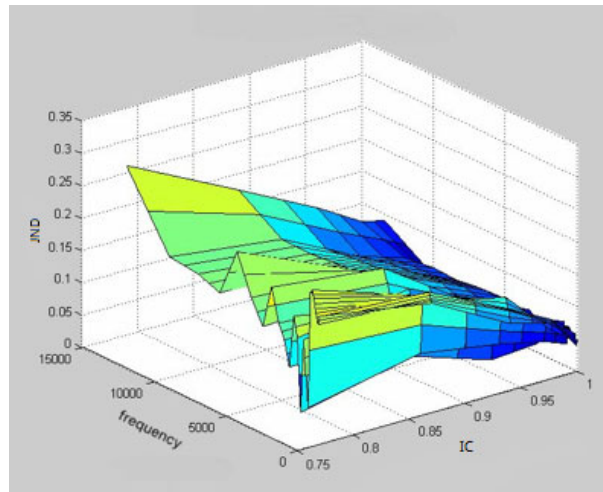

Fig. 5. Three dimensional curved surface of JND-Frequency-IC of interaural correlation

From Fig. 5, we can see that:

With the decrease of IC value (i.e. the increase of the width of the sound image), the JND is increasing, and the trend is obvious in the lower IC value. This conclusion is consistent with previous studies on the perceived characteristics of interaural correlation, that is to say, with the change of IC from 1 to 0, the value of JND is increasing, and the sensitivity of human ear to the sound image width decreases. When the IC value is close to 1, the surface is smooth and human ear is more stable in the perceptive ability of the sound image width. When IC values gradually change to 0, JND values fluctuate a lot, and human ear has an unstable perception of sound image width.

When the frequency of the sound signal is 50~500Hz, the JND value of IC does not change obviously. When the frequency is between 500Hz and 1000Hz, the JND decreases first and then increases, and then it stays stable when the IC value is larger, and it fluctuates greatly when the IC value is small. When the frequency exceeds 10 kHz, the value of JND increases dramatically, and it is difficult for the human ear to perceive changes in the sound image width.

3.4 JND Function Approximation Curve

In this paper, the linear least squares method is used to approximate the JND data of the IC, using the 6th order polynomial to represent the JND curve, the result is obtained by Matlab as follows:

(1) When the IC value is 1, the function expression of the JND curve of the perceived characteristic

curve is:

$$JND_1(n) = 1.1485 \times 10^{-7} n^6 - 7.6417 \times 10^{-6} n^5 + 1.8972 \times 10^{-4} n^4$$
$$-2.1 \times 10^{-3} n^3 + 1.06 \times 10^{-2} n^2 - 1.67 \times 10^{-2} n + 0.0478 \tag{7}$$

Root mean square error (RMSE) is:

$$RMSE(JND) = \sqrt{\frac{\sum_{n=1}^{24}(JND_1(n) - JND_n)^2}{24}} = 0.0081 \tag{8}$$

Where n represents the band number. Fig. 6 shows the relationship between the original value point and approximation curve of JND:
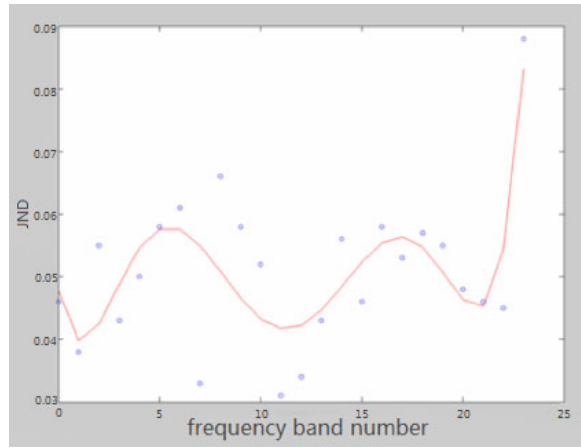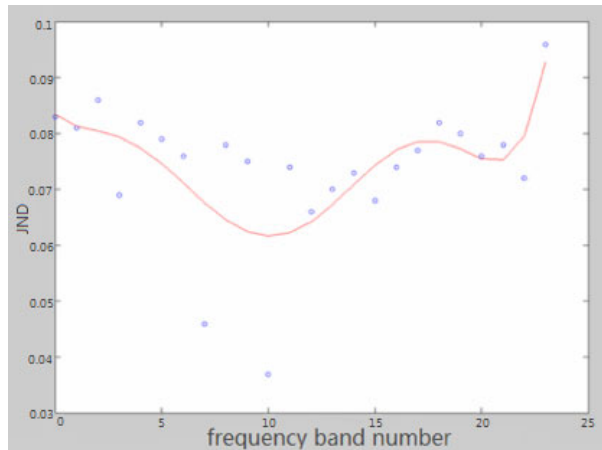

Fig. 6. The JND approximation curve when IC value is 1

(2) When the IC value is 0.98, the function expression of the JND curve of the perceived characteristic curve is:

$$JND_{0.98}(n) = 4.2961 \times 10^{-8} n^6 - 2.6523 \times 10^{-6} n^5 + 5.8345 \times 10^{-5} n^4 -$$
$$5.3625 \times 10^{-4} n^3 + 1.9 \times 10^{-3} n^2 - 3.6 \times 10^{-3} n + 0.0835 \tag{9}$$

RMSE is:

$$RMSE(JND) = \sqrt{\frac{\sum_{n=1}^{24}(JND_{0.98}(n) - JND_n)^2}{24}} = 0.0089 \tag{10}$$

Where n represents the band number. Fig. 7 shows the relationship between the original value point and approximation curve of JND:

Fig. 7. The JND approximation curve when IC value is 0.98

(3) When the IC value is 0.94, the function expression of the JND curve of the perceived characteristic curve is:

$$JND_{0.94}(n) = 4.4904 \times 10^{-8} n^6 - 1.9672 \times 10^{-6} n^5 + 1.1024 \times 10^{-5} n^4 + \\ 5.4321 \times 10^{-4} n^3 - 8.1 \times 10^{-3} n^2 + 2.81 \times 10^{-2} n + 0.1012 \tag{11}$$

RMSE is:

$$RMSE(JND) = \sqrt{\frac{\sum_{n=1}^{24}(JND_{0.94}(n) - JND_n)^2}{24}} = 0.0123 \tag{12}$$

Where n represents the band number. Fig. 8 shows the relationship between the original value point and approximation curve of JND:
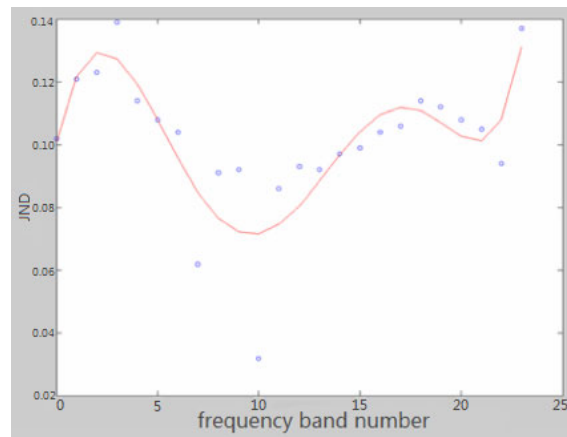

Fig. 8. The JND approximation curve when IC value is 0.94

(4) When the IC value is 0.87, the function expression of the JND curve of the perceived characteristic curve is:

$$JND_{0.87}(n) = 8.3503 \times 10^{-8} n^6 - 4.8398 \times 10^{-6} n^5 + 9.3706 \times 10^{-5} n^4 - \\ 5.9504 \times 10^{-4} n^3 - 6.9158 \times 10^{-4} n^2 + 8.6 \times 10^{-3} n + 0.168 \tag{13}$$

RMSE is:

$$RMSE(JND) = \sqrt{\frac{\sum_{n=1}^{24}(JND_{0.87}(n) - JND_n)^2}{24}} = 0.0242 \tag{14}$$

Where n represents the band number. Fig. 9 shows the relationship between the original value point and approximation curve of JND:
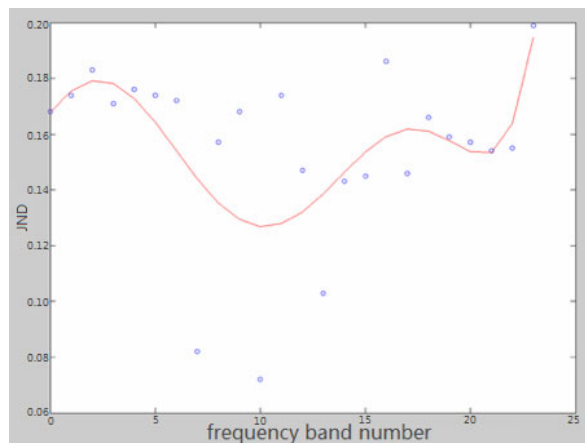

Fig. 9. The JND approximation curve when IC value is 0.87

(5) When the IC value is 0.77, the function expression of the JND curve of the perceived characteristic curve is:

$$JND_{0.77}(n) = 2.6652 \times 10^{-7} n^6 - 1.6275 \times 10^{-5} n^5 + 3.4869 \times 10^{-4} n^4 -$$
$$2.9 \times 10^{-3} n^3 + 6.5 \times 10^{-3} n^2 + 7.9 \times 10^{-3} n + 0.1851 \tag{15}$$

RMSE is:

$$RMSE(JND) = \sqrt{\frac{\sum_{n=1}^{24}(JND_{0.77}(n) - JND_n)^2}{24}} = 0.0394 \tag{16}$$

Where n represents the band number. Fig. 10 shows the relationship between the original value point and approximation curve of JND:
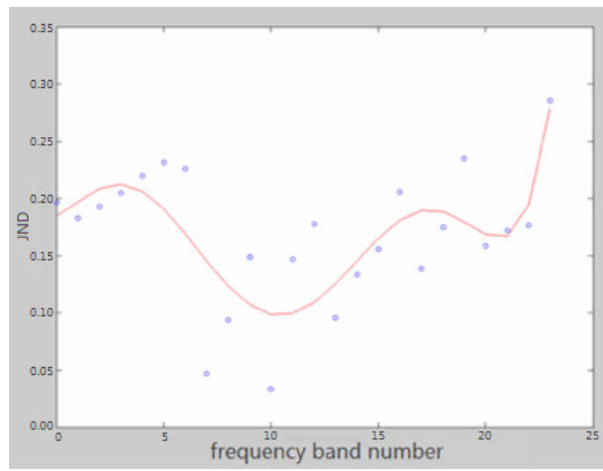

Fig. 10. The JND approximation curve when IC value is 0.77

4   Conclusion

To study the relationship between interaural correlation IC, frequency and JND, in this paper, the JND values of 24 bands were measured and analyzed under 5 different IC values, and the experimental data were processed by linear interpolation method, and the sample size was expanded. The experimental results show that the value of JND changes significantly with frequency, and the minimum value are found at 840Hz and 1370Hz. With the increase of IC, the JND of IC also increases correspondingly. Finally, a three-dimensional surface map of JND-frequency-IC is obtained, which can analyze the experimental data and explore the human ear perception mechanism, so as to provide further theoretical support for the efficient compression of 3D audio. In the later application of IC parameters, different bands can be treated differently. If the selection of testers and the selection of IC values are more comprehensive in subsequent studies, the theoretical knowledge of spatial audio coding can be further improved.

Acknowledgements

References

[1] C. Kim, R. Mason, T. Brookes, Initial investigation of signal capture techniques for objective measurement of spatial impression considering head movement, in: Proc. of the 124th Audio Engineering Society Convention, 2008.

[2] I. Pollack, W.J. Trittipoe, Binuural listening and Interaural noise cross correlation, The Journal of the Acoustical Society of America 31(1)(1959) 1250-1252.

[3] K.J. Gabriel, H.S. Colburn, Interaural correlation discrimination: I. Bandwidth and level dependence, The Journal of the Acoustical Society of America 69(5)(1981) 1394-1401.

[4] T.J. Cox, W.J. Davies, Y.W. Lam, The sensitivity of listeners to early sound field changes in auditoria, The Journal of the Acustica Society of America 79(1)(1993) 27-41.

[5] T. Okano, Judgments of noticeable differences in sound fields of concert halls caused by intensity variations in early reflections, The Journal of the Acoustical Society of America 111(1)(2002) 17-229.

[6] H. Lüddemann, H. Riedel, B. Kollmeier, Electrophysiological and psychophysical asymmetries in sensitivity to interaural correlation steps, Hearing Research 256(1-2)(2009) 39-57.

[7] S. Cao, R.M. Hu, Y.X. Peng, Research on spatial perceptible information estimation and it's application in spatial audio coding, [dissertation] Wuhan University, 2013.

[8] R. Mcaulay, M. Malpass, Speech enhancement using a soft-decision noise suppression filter, The Journal of the Acoustical Society of America 65(2)(1980) 137-145.

[9] J. Blauert, R.A. Butler, Spatial hearing: The psychophysics of human sound localization by Jens Blauert, Journal of the Acoustical Society of America 77(1)(1985) 334-335.

[10] J.C. Middlebrooks, D.M. Green, Sound localization by human listeners, Annual Review of Psychology 42(1)(1991) 135-159.

[11] Y.E. Ting, S.N. Zhu, new variable step-size LMS adaptive filtering algorithm and its performance in time delay estimation, Telecommunication Engineering 54(6)(2014) 776-777.

[12] R. Hu, H. Wang, W. Tu, C. Zhang, S. Wang, Relative contribution of frequency and time difference to selectivity for interaural time difference, Acta Acustica 39(6)(2014) 752-756.