Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 6244225, 11 pages
https://doi.org/10.1155/2022/6244225

Research Article

Multi-Tone Piano Transcription Based on Visualization and Knowledge Recognition Algorithm

Yingjing Fan

Chongqing Normal University College of Music, Chongqing 401331, China

Correspondence should be addressed to Yingjing Fan; 20131022@cqnu.edu.cn

Received 12 April 2022; Revised 1 June 2022; Accepted 11 June 2022; Published 30 July 2022

Academic Editor: Leipo Liu

Scientific computing visualization is also known as visualization. In an academic sense, visualization is a way of computing. It can convert data into intuitive geometric images, making it easier to understand. Visualization is the use of computer graphics and image processing techniques. It is a technical means to convert piles of seemingly disorganized data into graphics or images through special computer methods. Recording is the process of recording sound signals on a medium. Specifically, it is the process of converting sound into electrical signals through microphones and amplifiers and recording them with different materials and techniques. Piano tuning is the adjustment of the strings so that they reach a certain high pitch. This paper aims to study the analysis and evaluation of multi-tone piano transcriptions based on scientific computing visualization and knowledge recognition algorithms. It expects to improve the quality of piano recording with the support of scientific visualization technology and knowledge recognition algorithm. This paper briefly explains the big data, and on this basis, the visual data mining technology is mentioned, and the current situation in the visual data mining is explained. A matrix calculation of the characteristic matrix of piano timbres was carried out to achieve the resynthesis of the 25th harmony in piano timbres electronically. The results of the trials in this article demonstrate that when the Gaussian number changes to 5 during the recording process, the difference that exists of the acoustic voltage intensity and the noise level phase is clearly diminished. When the Gauss figure is 10, there is an inversion of the value of the sound pressure phase.

1. Introduction

Innovative emerging offerings such as the IoT and digital web have led to an extraordinary rise in the type and size of data in civil societies. Due to the limitation of the level of technology, there is a lot of information that can only be processed in batches. However, the continuous promotion of Internet technology makes more and more data information to be processed, and the original data processing and display methods are far from meeting the requirements. Based on this situation, scientific computing visualization came into being. Visualization technology needs to end up with users in the process of use, and it should not be overprocessed. When it can be graphical, use graphics and sketches as much as possible. The piano is a percussion musical device with a very broad tonal spectrum. Not just can pianos be enjoyed live, they can also be preserved using tape techniques. There are various types of audio taping methods, but in search of the optimum method of audio tapping, depending on the tonal features, the most fitting approach has always been pursued.

Visualization technology combines the power of man and machine, greatly improving people's work efficiency. Piano recordings recreate the power of listening to a track up close, reproducing the atmosphere and tone of a piece of music. The appreciation of art forms like piano can be enhanced to a degree. But not everyone has that.

This article innovatively uses scientific visualization techniques to analyze piano recordings. In addition, the method of sound spectrum recognition analysis and sound field perception analysis is also used in this paper. It collects and analyzes the piano sound characteristic matrix, innovates the piano recording method, and makes the recorded sound effect higher.

2. Related Work

With the development of brain imaging technology, brain science is gradually unraveling the mystery of human emotions. In this wave, data visualization has become the signature technology. Chen et al. reviewed recent literature in the field of connectomics and visualization on visualization of brain networks. They were particularly focused on techniques for visualizing brain networks at the macroscopic level. Compared to neuronal connections at the microscopic level, it reveals the structural and functional connectivity of the entire brain [1]. Sigtia et al. presented a model of monitored neighbourhood nets for use in making musical transmissions. Their framework proposed a military model consisting of an acoustical model and a linguistic product line model. The sonic market segment is a neighbourhood-wide system for approximating the probability of pitch in a frame of sound. The linguistic module is a retrospective aural structure that simulates pitch changes [2]. Ewert and Sandler explored the influence of instrument categories and used information from the scenes in question on the quality of the recordings. They developed a novel signal model with a variable-length spectrum, tailored for pitched percussion instruments such as piano [3]. For decades, sine polyphony and synthetic polyphony have been used to study the phenomenon of subjectively amplified octaves. Jaatinen et al. elaborated on this topic in a continuous-tone listening experiment with real orchestral samples. They used the stretching curves to analyze the collected data, applying an additive model in the manner of the Railsback curves. Although it was observed that the tuning curves of the orchestra modeling other curves differed, they were similar in quality. The experiments showed that the musical background had no significant effect on the results [4]. Yu et al. proposed a self-consistent approach. Their proposed sonification method uses the normal pattern of amino acid components of proteins to calculate vibrations. This approach to transcendence makes sure that the corresponding frequencies of oscillations within each molecule and between molecules are preserved [5]. Kitamura et al. studied the musical abilities of children with WS under participation. They presented two piano tones in sequence and asked children to respond nonverbally to judge the change in the second tone. The results showed that children with WS performed below the level predicted by their age (CA) in pitch recognition [6]. Togootogtokh et al. proposed a gesture tracking recognition method. For hand tracking, they proposed a model-based left-hand tracking and an appearance-based right-hand tracking technique. For virtual piano, they introduced a neural network (NN) approach to detect special gestures [7]. Johnson et al. presented a study of hand posture for pianists. In order to automatically assess the student's hand posture, they proposed a system. The system is able to recognize three types of postures. To train an artificial segmentation model, they tested two types of descriptors for functions: background image-based features and strength background features, which depict the contrast in the context of a given pupil's vicinity. The experiments demonstrated the validity of this approach. Of these, the height profile functioned best for the histogram and common line vegetation of the depth background feature [8]. Schwartz et al. investigated the effects of processing fluency on written musical metamemory. In Experiment 1, pianists created congruent and incongruent conditions by playing short sequences labeled with treble or bass clefs in the left or right hand, respectively, on a silent keyboard. In Experiment 2, judgments of learning (JOL) were tracked for each study sequence to measure prospective metamemory. In the consistency condition, JOL was higher, but recognition was not affected [9]. Sacha et al. presented a scheme that combines automated algorithms and interactive visualization. The framework identifies key scenarios for combining ML methods with interactive visualization through example illustrations, further elaborating the interactive ML process [10]. While these doctrines go some way to explaining visualization and logging skills for scientific computing, the union of the 2 is not readily apparent or pragmatic.

3. Multi-Tone Piano Transcription Analysis and Evaluation Method Based on Scientific Computing Visualization and Knowledge Recognition Algorithm

3.1. Overview of Knowledge Recognition. The meaning of knowledge has its own interpretation in each discipline. Therefore, no fixed explanation has been formed so far [11]. For example, the field of economics can be explained from an economic perspective, and philosophy can be explained from a dialectical perspective. Each discipline has its own point of view [12]. However, the traditional concept generally holds that knowledge is ordinary education, which has nothing to do with people's work and survival. Knowledge focuses on human development and has nothing to do with practicality. In modern dictionaries, knowledge is defined as the summary of knowledge and experience gained by people in the practice of transforming the objective world [13, 14]. Some scholars also define knowledge as "discipline," which refers to specialized knowledge or insight and experience that humans consider to be correct and true. Although knowledge has similarities, the forms and purposes of knowledge presentation are not the same. Knowledge can be divided into factual knowledge, principle knowledge, skill knowledge, and empirical knowledge [15]. Figure 1 shows a schematic diagram of the knowledge structure system.

To successfully utilize the value of knowledge, it is necessary to understand the basic characteristics of knowledge. It is generally believed that the characteristics of knowledge mainly include transferability, hierarchical utilization rate and return on investment, accelerated aging speed of knowledge update, and uncertainty value of knowledge. The transferability of knowledge is reflected in the transfer of knowledge within the organization. Among them, explicit knowledge is transferred through

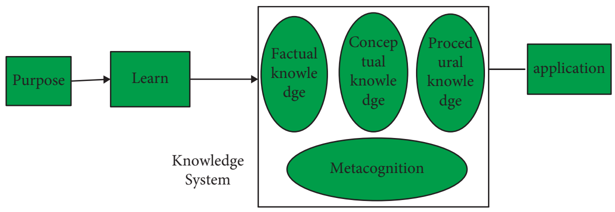


Figure 1: Knowledge structure system.

communication, while invisible knowledge needs to be transformed for a period of time to achieve transfer. Tacit knowledge is deeply engrained individual content that is resistant to formality or exchange. It normally resides in the shape of personally held experiences, images, senses, tacit understanding of groups, and skills in terms of technology, rather than clear knowledge expressed in words, language, images, etc. Tacit knowledge is very subjective, it cannot be explained by skill and perception, and it is not easy to document, record, transmit, and explain. It learns and accumulates in practical experience, hands-on practice, and continuous experimentation. Explicit knowledge can be encoded and presented, can be clearly explained, and is more objective. It can be systematically disseminated through coding using formal text, diagrams, etc. Hierarchy of knowledge refers to a partial description of events, not a complete representation of the entire time. Knowledge has a high utilization rate, and a return on income means that benefits are obtained from the use of knowledge. It can also continuously enrich the knowledge network and enhance the value of knowledge. In short, the more the knowledge is used, the more the value it adds. The acceleration of knowledge update rate and the acceleration of knowledge aging speed refer to the emergence of many branches in the contemporary era of accelerating knowledge generation, which complicates the overall knowledge structure. The core knowledge learned today is the introductory knowledge tomorrow. The uncertainty value of knowledge refers to the uncertainty of knowledge investment. There is a high degree of variation, and the value created by different people using the same knowledge is different.

3.2. Visual Data Mining Technology. Data visualization is a process of using statistical analysis techniques to retrieve key insights out of large datasets and then display them in a visual format.

According to the basic idea of the classical multi-dimensional scaling method, we can roughly list the distance matrix function expression:

$$\left(a_c - a_h\right)'\left(a_c - a_h\right) = a_c' a_c + a_h' a_h - 2a_c' a_h. \tag{1}$$

We denote the distance of the matrix as l, so a is a simple composition of l.

$$\varphi = \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_\mu}{\alpha_1 + \alpha_2 + \cdots + \alpha_p} \geq \varphi_\gamma, \tag{2}$$

where φ_γ represents the given variation contribution rate, and the value range of φ is uncertain in the whole calculation process.

A decision-making tree is basically a decision classification and profiling method for performing sorting and quality analysis on data. The exact structure is shown in Figure 2.

The principle process of the ID3 algorithm is to find the roots and then find the optimal solution from the results obtained.

$$Q(A) = -\sum_{k}^{l} g\left(a_k\right)\log_2 F\left(a_k\right). \tag{3}$$

When $g(a_1) = g(a_2)$, $Q(A) = 1$.

$$j(f, p) = -\frac{f}{u+p}\log_2\frac{f}{u+p} - \frac{p}{u+p}\log_2\frac{p}{u+p}. \tag{4}$$

Equation (4) shows the information content required for the proper clustering of the tree.

$$Y(B) = \sum_{K}^{s} \frac{f_k + l_k}{f + l} j\left(f_k, l_k\right), \tag{5}$$

where B relates to a single child of the FC tree, j refers to the average value of expectations, and $Y(B)$ refers to the total value of beliefs.

$$\text{Info}(S) = -\sum_{1}^{O} U_a \log_3\left(U_a\right), \tag{6}$$

where U_a stands for the sample as a share of the total human race.

$$\text{Info}_x(S) = \sum_{1}^{o} \frac{\left|S_1\right|}{|S|} * \text{Info}(S), \tag{7}$$

where X stands for characteristics of the property of the model and $\text{Info}_x(S)$ stands for the desired message.

$$\text{Gaint}(S) = \text{Info}(S) - \text{Info}_X(S). \tag{8}$$

Equation (8) denotes the discrepancy that exists among new and older demand for knowledge.

$$W_a = \alpha + \beta W_{a-1} + \delta_a, \tag{9}$$

where W_a stands for needs and β stands for variation between needs, while α stands for a constant.

$$\overline{W}_a = \frac{\sum_s^{a-1} W_a}{s}, \tag{10}$$

$$F_a^3 = \frac{\sum_j^{a-1}\left(W_j - \overline{W}_j\right)^3}{j-2}. \tag{11}$$

Equations (11) and (12) stand for the forecast in terms of cost estimates.

$$\text{New\_}R\left(U, C_s\right) = \sum_{c} \frac{\left|U_l\right|}{|U|} R\left(U_l\right). \tag{12}$$

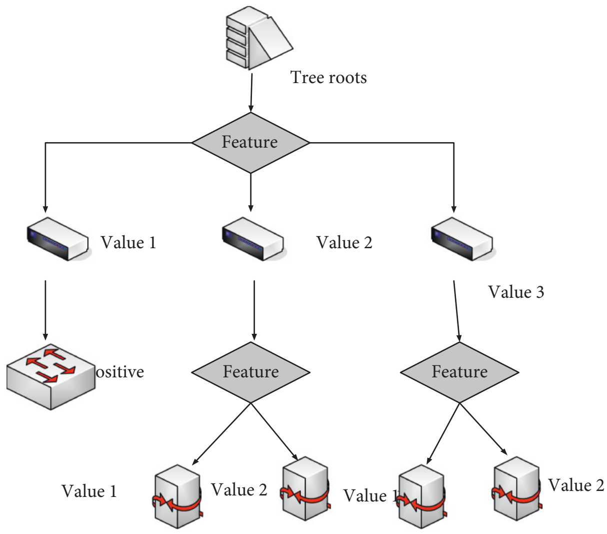


Figure 2: Decision tree structure.

Equation (13) stands for a representation of the sub-tree entropy of info as a function, and U_l represents the sample attribute feature.

$$G(Q) = H(a,b) - Y(Q), \tag{13}$$

where $G(Q)$ represents the information gain of attribute B.

$$K(R) = -\sum D_A \log_2 (D_A), \tag{14}$$

where k stands for the child set and D stands for the level of detail in the child set.

The classical multi-dimensional scaling method can simplify the multi-dimensional attribute variables, and then the data can be limited and classified. A similar technique is the intelligent distribution algorithm. With the continuous development of multi-intelligence western distribution algorithms, researchers try to apply the theory to practice. They also believe that the gradient projection method can be used to design distributed optimization algorithms. Figure 3 shows a schematic diagram of the distributed optimization structure.

Multi-agent can be divided into directed graph and undirected graph. The function expression of a directed graph can be expressed as

$$C_{hf} = \begin{cases} 1, & \\ 0, & \end{cases} \quad (h,f) \in \alpha, \tag{15}$$

where C_{hf} represents the weight of the directed edge and h, f represents the directed node.

$$k_{hl}(h) = \sum_{f=1}^{l} C_{hf}, k_{\text{out}}(h) = \sum_{f=1}^{l} C_{hf}. \tag{16}$$

The in-degree matrix is $K_{hl} = \lfloor k_{hf} \rfloor$, and the out-degree matrix is $K_{\text{out}} = \lfloor k_{fh} \rfloor$. When $K_{hl} = K_{\text{out}}$, we consider it to be a weighted smooth map.

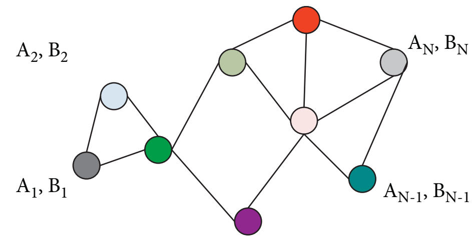


Figure 3: Distributed optimization structure.

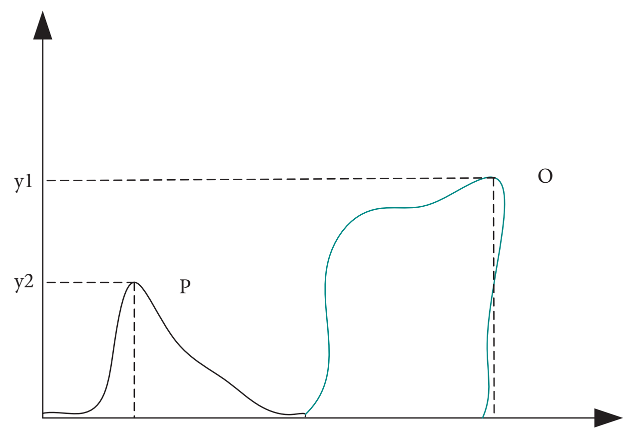


Figure 4: Joint matrix image.

$$g_{hf} = \begin{cases} -c_{hf}, & h \neq f, \\ k_{hm}(h), & h = f, \end{cases} \tag{17}$$

where $g(hf)$ represents a directed graph matrix. The specific results are shown in Figure 4.

$Y = (y_{cx})$ represents a continuous uniform Markov process transition rate matrix. We can generalize it with the following function expression:

$$B\{\alpha_{v+z} = d|\alpha_v = c\} = \begin{cases} 1 + y_{cx}z + f(z), & c = z, \\ y_{cx}z + f(z), & c \neq z, \end{cases} \tag{18}$$

where $y_{cx} = -\sum_{c \neq x} y_{cx}$, $y_{cx} \geq 0$, $\lim_{z \longrightarrow 0} f(z)/z = 0$. For a discrete Markov process, its transition probability matrix can be expressed as

$$g_{cx} = B\{\alpha_{v+z} = d|\alpha_v = c\}, \quad v \in M, \tag{19}$$

where g_{cx} represents Markov's transition probability matrix.

3.3. Piano Overview. Sound is a familiar entity, and the capture of sound calls for the use of audio transcription media. Audio recording came into the general public eye with the advent of the gramophone.

The piano has 88 keys, 52 of which are white and 36 are black. The 88th key, c5, is the highest note, and the fundamental frequency of the sound is 4186.01 Hz. Piano strings are made of different materials and have different pronunciations. In terms of general build, the piano consists of three main pieces: the case, the mechanism, and the tone generator. The motion mechanism consists of the hallmark, linkage, butter, and other components. The sound mechanism is made up of three parts: the nut, the bellows, and the fabric bracing. In a finished composition of music, not too many vibrational waves of varying frequencies are present, which we call in academic terms rhythm. The twelve mean meters determine the basic frequency of each note of an organ. In music criticism, two notes with the same name differ in pitch by one degree of separation. It is well known that piano keys consist of white keys and black keys. However, in both white and black keys, there is a chromatic difference between the two adjacent keys. If two white keys are joined in an arrangement, the distance between the white key and the intervening black key is a semitone if the two white keys have a black key in them. Figure 5 shows the tonal structure of the piano height.

Length means the length of time it takes to produce a tone. It is essentially the time of vibration of the string. The identification of the length of a sound is therefore intended to achieve the identification of the time of the string's vibrating sound signature. In the pursuit of musical aesthetics, although the duration of a tone cannot be altered, its lasting time can be altered. Figure 6 shows a graphic representation of the musical generation scale.

When we use the same intensity and location to play the piano, the ringing may not be identical. This is because the piano is influenced not only by the amount of force but also by the strength of the different strings due to their respective qualities and resonance channels. The piano is known as the "king of playing tools." The bass is deep and thick, the midrange natural, and the treble bright. With its distinctive appeal, it is the musicians' favourite instrument and the most versatile instrument. Indeed, with brief practice, it is possible to distinguish between the different tones.

4. Multi-Tone Piano Transcription Analysis and Evaluation Experiment Based on Scientific Computing Visualization and Knowledge Recognition Algorithm

4.1. Data Collection. In this lab, we sort and define the diverse categories of piano music. Each genre consists of music from the same time. These music clips were mainly taken from mobile phone audio. It is as follows.

Based on the data in Table 1, experimental subjects of category music, Hip Hop, Blues, and Rock music are used and then be divided into four groups. However, there is a difference in the number of training and validation as well as the amount of trials. The count of training was revised to 320, the count of validation to 50, and the count of trials also to 50 for comparison.

Based on the figures in Table 2, we divided the experiments into four sets with Classical, Hip Hop, Blues, and Rock music. The quantity of training was 350, the quantity of validation was 60, and the quantity of testing was also 60.

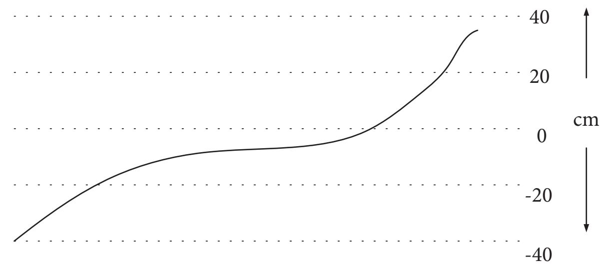


Figure 5: Piano pitch tuning curve.

4.2. Twelve Mean Laws. The twelve laws use the rule that the proportion of equal frequency of opposite semitones is equal. This makes what used to be equal temperamentally in pitch a true homophony in the present sense. It also achieves the consistency of intervals, chords, and the same mode relationship of various keys. Due to the proportional relationship of the twelve equal laws, it allows for the transfer of a variety of intricate creations to realize the diverse aesthetic needs of the listener.

According to the data in Table 3, if an interval is shifted to a pitch range above it, the rhythm of the rhythm of the beat rises. On the contrary, if the spacing is displaced to a lower sphere, the beat's rhythm is reduced.

4.3. Pure Piano Temperament. Although now in use, the dodecaphonic piano was not in use when the dodecaphonic piano first came into existence.

According to the data in Table 4, the main strength of pure meter is the large variety of wolfish fifths, but there are also weaknesses. There are many wolf fifths in pure meter and a large variety of minor thirds, making the whole pure meter system too complex. The distribution range of the impure fifth is wider, which makes the transfer more difficult.

5. Multi-Tone Piano Transcription Analysis and Evaluation Analysis Based on Scientific Computing Visualization and Knowledge Recognition Algorithm

5.1. Kurtosis Matrix Analysis. The kurtosis is also known as the kurtosis coefficient. Intuitively, kurtosis reflects the sharpness of the peak. A high skewness means that the added squareness is due to low level poles that are either greater or less than the average.

According to the data in Figure 7, when the hold period is 20 s, the peaks of all four sets are low and the kurtosis value of all four sets is 50. When the hold period is 40 s, the peaks of the first set of data reach 450, the kurtosis of the second set of data is 20, the peaks of the third set of data is 50, and the kurtosis of the fourth set of data is 50. It can be seen that, except for the first set, the other three sets are to remain in the trough. When the hold period is 60 s, the peak of the first

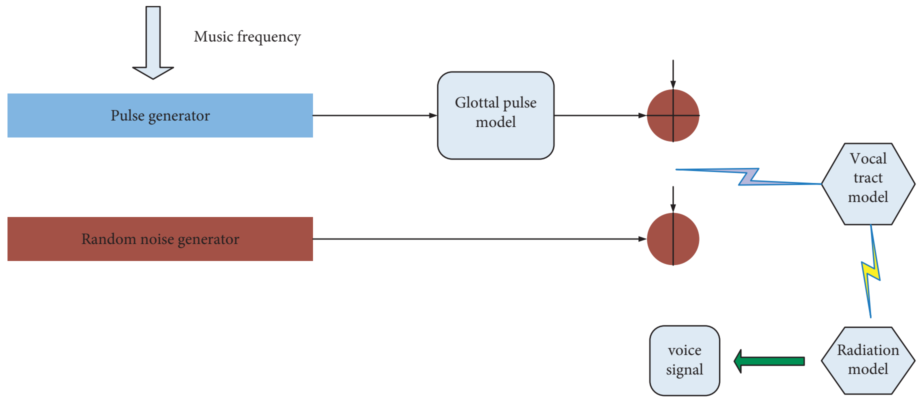


Figure 6: Music production model.

Table 1: Experimental dataset.

Group	0	1	2	3
Style	Classical	Hip Hop	Blues	Rock
Number of training	320	320	320	320
Number of verifications	50	50	50	50
Number of tests	50	50	50	50

Table 2: Experimental music data.

SIG	0	1	2	3
Mode	Classical	Hip Hop	Blues	Rock
Number of training	350	350	350	350
Number of verifications	60	60	60	60
Number of tests	60	60	60	60

Table 3: Interval coordination.

Interval	Key tone	Pitch frequency	Overtone frequency	—	—
Four degrees	31E	163	499	1150	1312
—	36A	220	440	1050	1320
—	Beat	—	×	√	√
Five degrees	32E	163	499	1200	1421
—	37A	250	440	1235	1478
—	Beat	—	×	√	×

Table 4: Pure law cycle.

The Bloc	Solid five degrees	Big three degrees	Lesser third
bb	0	0	7.23
f	0	0	7.21
g	−22	0	22
a	0	0	0
b	22	43	22
e	0	0	0
c	0	0	0

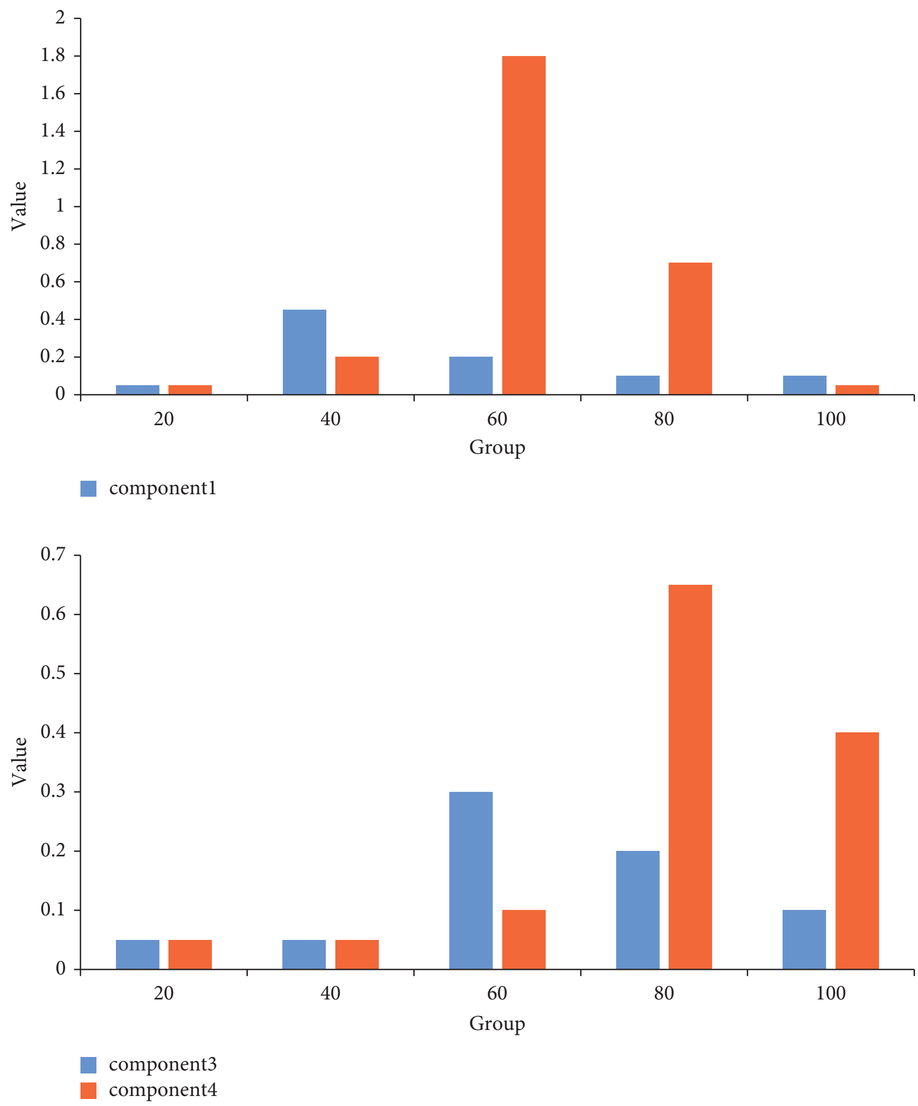
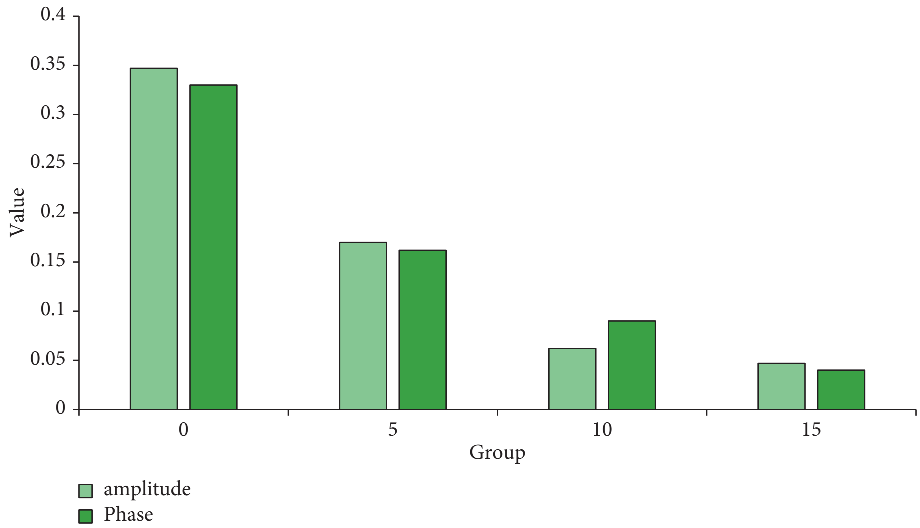


Figure 7: Original chromatographic curve.

set of data is 20, the peak of the second set of data is up to 1800, the peak of the third set of data is up to 300, and the peak of the fourth set of data is 100. When the hold period is 80 s, the peak of the first set of data is 100, the peak of the second set of data is up to 700, the peak of the third set of data is 200, and the peak of the fourth set of data is up to 650. When the hold period is 100 s, the peaks are 100 for the first set of data, 50 for the second set of data, and 100 for the third set of data, and the kurtosis of the fourth set of data is 400. The four groups of kurtosis showed a decreasing trend. According to these data, even if the same music is intercepted, if the interception time is different, the kurtosis will change differently.

5.2. Sound Field Perception Analysis. Sound mixing is a domain that brings together sound and science and technology. This way of recording focuses on collecting varying genres of sound by adapting the range from a sound source to the distance between the recording mike. There is better ambient spatiality for distant audio recording, but ambient bounce is also recorded and is more destructive to the music as a whole. Closer recording captures a more textured sound with less noise from the environment.

The A-B recording method consists of placing two microphones on top of each other for audio taping. It causes the cones of both microphones to be oriented towards the vocalist. The directionality of the microphones is generally used in an omnidirectional recording method. In the A-B system, there is a distance between the two microphones, so the sound has a certain intensity and time difference, and the resulting audio is more comparable in terms of spatial context to that heard by the person's voice, and the sound is more aesthetically pleasing.



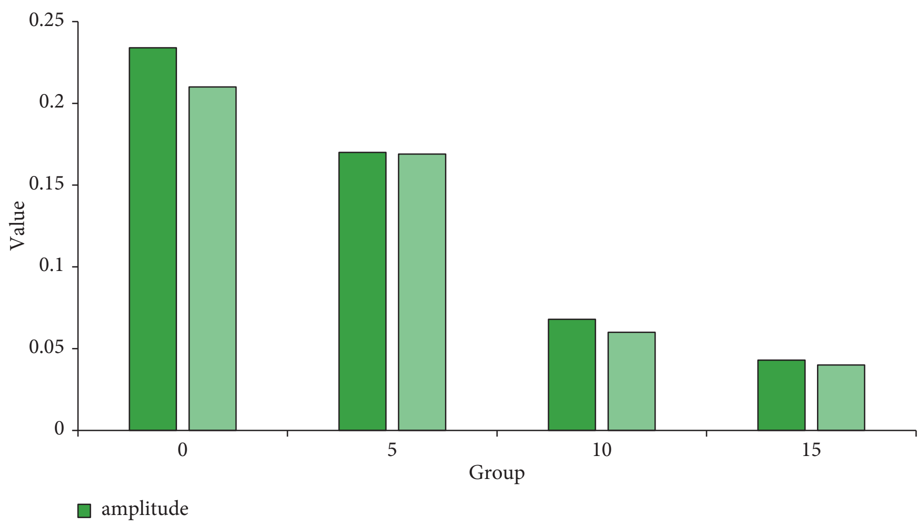


Figure 8: Sound field perception data.

During the recording process, the recording on each track is checked. It is necessary to replace redundant sounds and bad parts with other recording backups that appear during the recording process and to delete the paused parts of the performance.

Based on the figures in Figure 8, the two parameters shown are the acoustic volume magnitude and the acoustic phase. When there is no interference during the taping process, there is a difference that exists between the SPL amplitude and the SPL phase. However, when it is 5, the gap

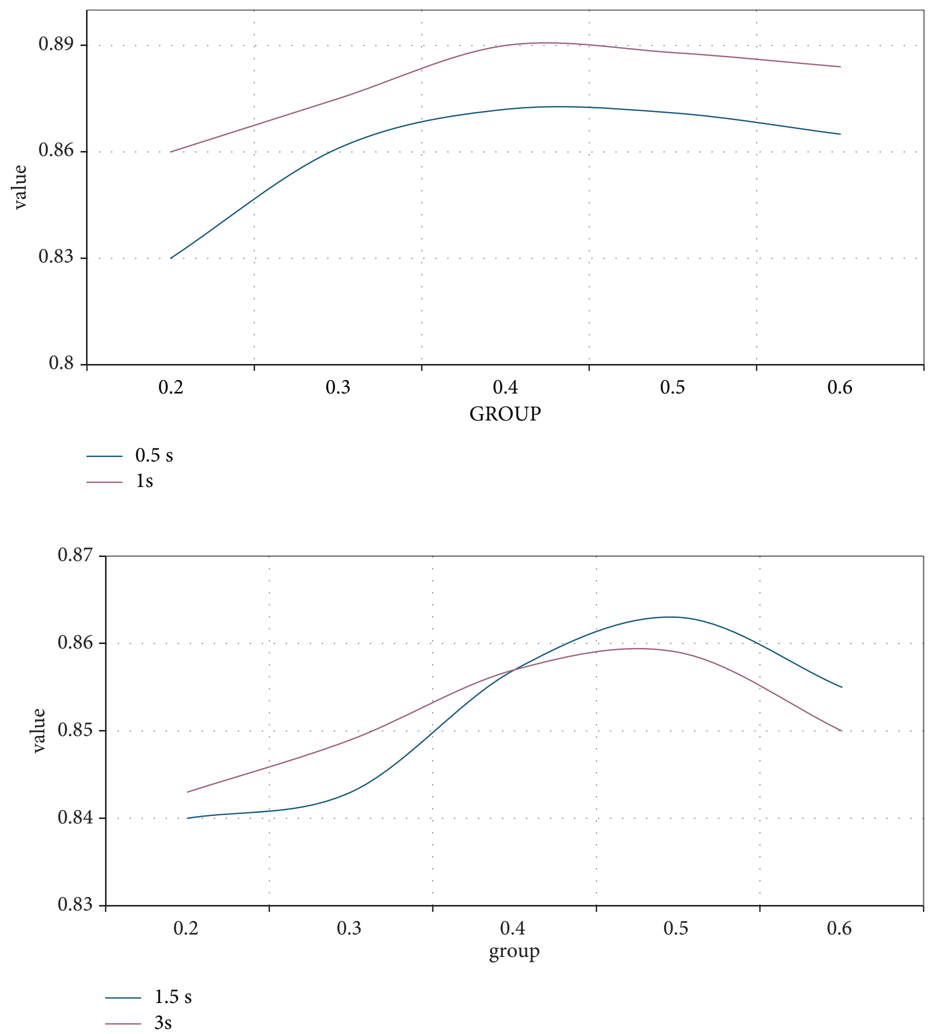


Figure 9: Micro-window analysis.

is significantly reduced. When it is 10, an inversion occurs. When it is 15, the gap is minimal. This indicates that this is the most ideal environment for recording.

5.3. Sound Spectrum Identification and Analysis. In our experiments, we classify a variety of types of piano pieces. The specific analysis is as follows.

Based on the data in Figure 9, we divided the time windows into four categories: 0.5 s, 1 s, 1.5 s, and 3 s. As the period of the analytical window gradually increased, the overall microscopic data at 0.5 s exhibited an upward and then downward trend. The trend at 1 s was similar to that at 0.5 s, but the downward trend was smaller than that at 0.5 s. Although the results at 1.5 s and 3 s showed a more similar trend, the overall magnitude was smaller.

As can be seen from the data in Figure 10, there are similarities in the experimental setup with the microscopic. It also divides the time window into four categories: 0.5 s, 1 s, 1.5 s, and 3 s. It can be seen from both sets of graphs that the graphs present the worst values when the time is 3 s, both from a micro-perspective and macro-perspective.

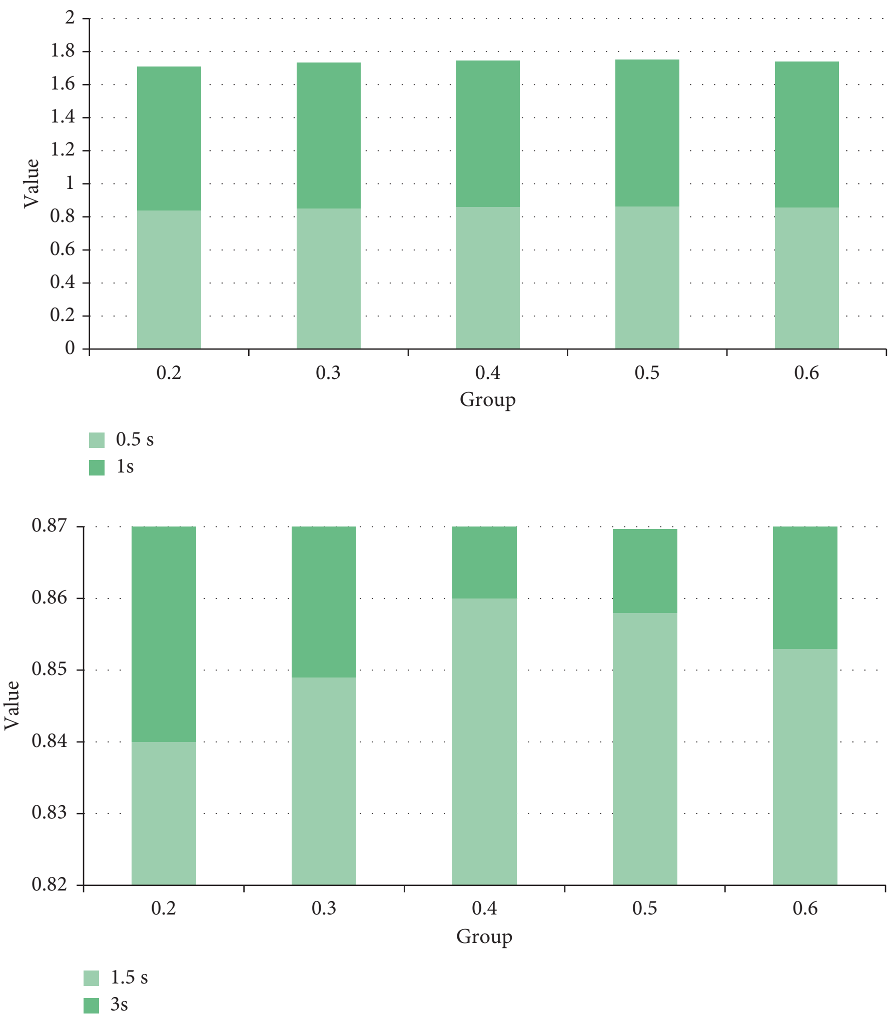


Figure 10: Macro-window analysis.

6. Conclusion

The current traditional computing power cannot meet the needs of production, so scientific visualization techniques have emerged as a current topic of interest. With the upgrading of the material world, the spiritual world is constantly changing in terms of needs. Piano music can nurture the spirit and is favoured by most people. The art of piano, however, has a high level of consumption and the question of how to reach out to the general public is one that needs to be addressed. The integration of high-performance computing and piano recital technology can enhance sound recording and satisfy higher awareness of aesthetic demands. Although some achievements have been made in this paper, there are still shortcomings. (1) The visualization research on traditional statistical application methods only introduces the multi-dimensional scaling method, which is too simple. (2) The data obtained in the piano music recording experiment may have certain limitations due to the defects of the equipment.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] W. Chen, L. Shi, and W. Chen, “A survey of macroscopic brain network visualization technology,” *Chinese Journal of Electronics*, vol. 27, no. 5, pp. 889–899, 2018.

[2] S. Sigtia, E. Benetos, and S. Dixon, “An end-to-end neural network for polyphonic music transcription,” *IEEE/ACM Transactions on Audio Speech & Language Processing*, vol. 24, no. 5, pp. 927–939, 2017.

[3] S. Ewert and M. Sandler, “Piano transcription in the studio using an extensible alternating directions framework,” *IEEE/*

ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 11, pp. 1983–1997, 2019.
[4] J. Jaatinen, J. Pätynen, and K. Alho, "Octave stretching phenomenon with complex tones of orchestral instruments," *Journal of the Acoustical Society of America*, vol. 146, no. 5, pp. 3203–3214, 2019.
[5] C.-H. Yu, Z. Qin, F. Martin-Martinez, and M. J. Buehler, "A self-consistent sonification method to translate amino acid sequences into musical compositions and application in protein design using artificial intelligence," *ACS Nano*, vol. 13, no. 7, pp. 7471–7482, 2019.
[6] Y. Kitamura, Y. Kita, Y. Okumura, and Kaga, "Discrepancy between musical ability and language skills in children with Williams syndrome," *Brain & Development*, vol. 42, no. 3, pp. 248–255, 2020.
[7] E. Togootogtokh, T. K. Shih, W. Kumara, S.-J. Wu, and Sun, "3D finger tracking and recognition image processing for real-time music playing with depth sensors," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9233–9248, 2018.
[8] D. Johnson, D. Damian, and G. Tzanetakis, "Detecting hand posture in piano playing using depth data," *Computer Music Journal*, vol. 43, no. 1, pp. 59–78, 2020.
[9] B. L. Schwartz, Z. F. Peynircioğlu, and J. R. Tatz, "Effect of processing fluency on metamemory for written music in piano players," *Psychology of Music*, vol. 48, no. 5, pp. 693–706, 2019.
[10] D. Sacha, M. Sedlmair, L. Zhang, and Lee, "What you see is what you can change: human-centered machine learning by interactive visualization," *Neurocomputing*, vol. 268, no. dec.11, pp. 164–175, 2017.
[11] P. Nickleson, "Transcription, recording, and authority in 'classic' minimalism," *Twentieth century Music*, vol. 14, no. 03, pp. 361–389, 2017.
[12] S. Karam, M. Nagahi, V. Dayarathna, J. Ma, and R. M. Jaradat, "Integrating systems thinking skills with multi-criteria decision-making technology to recruit employee candidates," *Expert Systems with Applications*, vol. 160, no. 113585, pp. 1–16, 2020.
[13] C. Leong, "Technology & recruiting 101: how it works and where it's going," *Strategic HR Review*, vol. 17, no. 1, pp. 50–52, 2018.
[14] S. Mainka, "Algorithm-based recruiting technology in the workplace," *Texas A&M Journal of Property Law*, vol. 5, no. 3, p. 8, 2019.
[15] A. T. Wynn and S. J. Correll, "Puncturing the pipeline: do technology companies alienate women in recruiting sessions?" *Social Studies of Science*, vol. 48, no. 1, pp. 149–164, 2018.