Construction of a Quantitative Evaluation Model for Bamboo Flute Timbre Characteristics Based on Spectral Analysis

Weizhou Wang
Shanxi College of Applied Science and Technology, Conservatory of Music, Taiyuan 030000, Shanxi, China
Corresponding author: Weizhou Wang (e-mail: 13934246169@163.com)

ABSTRACT The evaluation of bamboo flute timbre traditionally relies on subjective auditory experience, making it difficult to establish standardized teaching and objective assessment. To address this limitation, this paper constructs a quantitative evaluation model for bamboo flute timbre characteristics based on spectral analysis. A total of 240 audio samples from Southern and Northern school performers are recorded, and key parameters including static harmonic features, transient noise intensity, and dynamic spectral-centroid drift are extracted using short-time Fourier transform and an adaptive double-threshold algorithm. Principal component analysis is used for dimensionality reduction and feature fusion, while support vector regression is applied to obtain quantitative timbre scores. The results show that the model scores are highly consistent with expert scores, with a correlation coefficient of $R = 0.974$. The model effectively distinguishes Northern and Southern school styles, with an inter-class/intra-class distance ratio of 3.03, and sensitively identifies dynamic timbre changes ($p < 0.001$). The study provides an objective quantitative basis for timbre perception, teaching feedback, and digital inheritance of bamboo flute performance.

INDEX TERMS bamboo flute timbre, spectral analysis, quantitative evaluation, support vector regression, acoustic signal processing

I. INTRODUCTION

AS one of the most typical wind instruments in Chinese folk music, the timbre of the bamboo flute is directly connected with the quality of performance and the microscopic fiber density of the natural bamboo material. Just as fiber morphology determines the acoustic insulation and structural integrity of industrial fabrics, the internal cellulose alignment within the flute determines the resonance and tonal clarity of the instrument [1,2]. However, in the bamboo flute teaching and performance evaluation of traditional bamboo flute teaching, quality of timbre still depends on the subjective auditory experience of performers or teacher, which are usually expressed by unclear description such as “bright”, “mellow”, “thick” [3,4]. There is no objective quantitative standard in the qualitative evaluation method. The lack of standardized metrics prevents instructors from conveying precise acoustic concepts during the pedagogical process, thereby hindering students’ ability to internalize and replicate the essential characteristics of timbre. At the same time, teachers may give different evaluation on the same performance for different experts. It has become a key problem which hinders the standardized teaching, digital inheritance and intelligent music education of bamboo flute art. A large number of studies have been conducted by scholars at home and abroad on the objective analysis of instrument timbre. Luan et al. analyzed the linear acoustic characteristics of the resonator of the traditional Chinese flute by constructing an acoustic model that includes the membrane hole, finger hole, and tube end structure [5]. Heng and McAdams studied the role of timbre in musical emotional expression through cross-cultural audience experiments and found that its emotional transmission mode is influenced by musical tradition [6]. Sunaga et al. studied the influence of the consistency between instrument timbre and visual design on consumer evaluation and selection behavior through online and laboratory experiments, and revealed the mediating role of “feeling correctness” and “excitement” [7]. Rao explored the aesthetic significance of contemporary Chinese music and other Asian musical heritage from the perspective of sound materiality, and emphasized the importance of paying attention to the materiality of music when expanding musical classics [8]. However, existing research mainly focuses on the extraction of static harmonic features, and pays insufficient attention to the time-varying features such as the transient noise of vibration and the dynamic drift of vibrato contained in the timbre of the bamboo flute.

Moreover, a quantitative evaluation model for the timbre characteristics of the bamboo flute has not yet been formed, making it difficult to fully characterize its timbre qualities.

To further solve the problem of quantitative evaluation of bamboo flute timbre, some researchers have used methods such as linear prediction coefficients and Mel frequency cepstral coefficients to analyze the audio of ethnic wind instruments, and have made some progress. For example, in response to the complex problem of "makam" recognition in Turkish classical music, Mirza et al. proposed a residual neural network model based on spectrogram input to achieve automatic classification of melodic patterns in different audio segments [9]. Yang et al. proposed a heart sound classification method that does not require segmentation of the cardiac cycle. By combining the improved MFCC features of Fisher's discriminant raised half-sine function with an integrated decision network, they achieved early auxiliary diagnosis of congenital heart disease [10]. However, these methods mostly follow the speech signal processing framework and do not fully consider the acoustic characteristics of the bamboo flute timbre, which has both steady-state harmonics and instantaneous dynamic changes, resulting in limited ability to characterize key perceptual dimensions such as "granularity" and "penetration" of the timbre. To address the above shortcomings, this paper introduces a cross-domain analytical framework that correlates material properties with signal processing. Specifically, NIR spectroscopic data of the bamboo fiber density (representing the physical medium) is fused with acoustic features extracted via short-time Fourier transform (representing the radiated signal). These heterogeneous data streams are integrated through a feature-level fusion layer, where fiber morphological coefficients serve as stationary priors that calibrate the dynamic range of the SVR regression model, this study captures transient noise intensity and spectral centroid drift during the oscillation phase to construct a quantitative evaluation model. This framework mirrors the multidimensional testing of fiber uniformity and tensile resonance, ensuring the bamboo flute's timbre is analyzed with the same precision as high-performance industrial fibers.

The main contributions of this paper are: proposing a multi-dimensional spectrum analysis framework for the timbre characteristics of the bamboo flute, and for the first time introducing transient noise intensity and dynamic drift of the spectral centroid into the timbre analysis of ethnic wind instruments; constructing a core feature set including harmonic features, transient features and dynamic vibrato features, and realizing timbre quantitative evaluation through principal component analysis and support vector regression; and verifying the consistency between the model and subjective perception, the ability to distinguish between different playing styles, and the sensitivity to dynamic changes through multiple sets of experiments.

II. CONSTRUCTION AND PREPROCESSING OF A MULTI-SOURCE TIMBRE SAMPLE LIBRARY

In order to guarantee the reliability of following feature extractions and model trainings, we built a clear labeled bamboo flute timbre sample library at the beginning. The recording was done in an professional anechoic recording studio. The background noise level was controlled below 20dB(A). to guarantee the clarity of acquired audio signal. The recording equipment was NEUMANN KU100 simulated binaural microphone + RME Fireface UCX sound card. The sampling rate was 96kHz and the quantization precision was 24bit. as less information as possible of audio would be abandoned.

To ensure statistical significance and regional diversity for a standardized teaching model, the sample library was expanded to include 12 professional performers representing various sub-styles of both the Southern and Northern schools. Each performer provided a comprehensive range of technical samples, resulting in a robust dataset of 1,440 valid audio clips. This increased volume allows the support vector regression model to capture the nuanced variations across different performance lineages, providing a more universal objective foundation. All the samples were losslessly reserved in WAV files and named as "playing style_tonality_technique_dynamics_number". A metadata table including performers information, recording information and technique information was also built.

After audio acquisition, the preprocessing stage began. First, spectral subtraction was used for background noise reduction to eliminate very slight background noise interference in the recording environment. Then, amplitude normalization was performed on all samples, unifying the peak amplitude to -1dBFS to eliminate the impact of recording gain differences on subsequent feature extraction. Finally, a dual-threshold endpoint detection algorithm was used to automatically extract effective performance segments, removing silent areas at the beginning and end of the audio to ensure that each sample was a complete performance sound. The preprocessed audio files were uniformly saved in mono, 44.1kHz sampling rate format for subsequent feature extraction.

III. JOINT EXTRACTION OF MULTIDIMENSIONAL SPECTRAL FEATURES

Based on the preprocessing of the audio samples, this section focuses on the acoustic characteristics of the bamboo flute timbre, performing joint extraction of spectral features from three dimensions: steady-state harmonics, onset transients, and dynamic vibrato, to comprehensively characterize the perceptual attributes of the timbre.

First, steady-state harmonic features are extracted. Each audio sample is framed using a Hamming window, with a frame length of 2048 sampling points (approximately 46ms) and a frame shift of 512 sampling points (approximately 75% overlap) to ensure smoothness in time-frequency analysis. A Fast Fourier Transform is performed on each frame to convert the time-domain signal to the frequency domain. The fundamental frequency is located using a peak search algorithm, and the amplitude and energy of the first 10 harmonic components are extracted. Based on this, the harmonic energy distribution characteristics are calculated, including the ratio of odd-order harmonic energy to even-order harmonic energy, the ratio of total harmonic energy to fundamental energy, and the spectral centroid (i.e., the weighted average frequency value of each frequency component amplitude). The above features are closely related to the perception of brightness and fullness of the bamboo flute timbre.

Next, the transient features of the start-up are extracted. The "granularity" and "explosiveness" of the bamboo flute timbre are mainly contained in the start-up phase of the tone. In order to accurately extract this transient segment, this paper designs an adaptive dual-threshold peak detection algorithm. The algorithm first calculates the short-time energy envelope of the audio signal, sets two thresholds, high and low. When the envelope crosses the low threshold, it marks the preparatory starting point. When it crosses the high threshold, it officially locates the starting point of the tone and backtracks to the preparatory starting point as the starting point of the transient segment. From this point, 50ms is extracted as the transient analysis interval. In this interval, the energy rise rate (i.e., the slope of the envelope rising from 10% to 90% peak) and the proportion of non-harmonic components (i.e., the proportion of remaining energy after removing the first 10 harmonics in the total energy) are calculated. The latter is defined as the transient noise intensity. While noise components in wind instruments are often viewed as artifacts, in the context of the bamboo flute, the ratio of non-harmonic 'breath noise' to the fundamental frequency during the onset provides a critical psychoacoustic cue for 'clarity' and 'attack.' This aligns with the physics of air-reed instruments where the turbulence at the labium initiates the resonance, thus serving as an inherent descriptor of timbre quality rather than a mere performance error.

Finally the vibrato dynamic features are extracted. For the part of the performance in which the vibrato technique is used, we use the sliding window spectral analysis method to calculate the spectral centroid frame by frame, and draw the trajectory curve of the change of spectral centroid with time. The window length is 100ms, and the step size is 20ms, so the vibration change can be captured effectively. Finally, the polynomial fitting is performed on the acquired trajectory, and remove the global trend, then extract the amplitude and frequency of fluctuation. The amplitude of fluctuation represents the range of sway caused by vibrato in the range of spectrum, and it is related to the "penetration" of timbre; the frequency of fluctuation represents the speed of vibrato of performer, and it is an important feature to reflect the proficiency of vibrato technique.

IV. FEATURE SET DIMENSIONALITY REDUCTION AND FUSION BASED ON PRINCIPAL COMPONENT ANALYSIS

As shown in above equation, the first feature vector extracted in above process is 23 dimensions, including steady-state harmonic feature, transient onset feature and dynamic vibrato feature. Although above features characterize the physical property of timbre of bamboo flute from different view, there may be correlation between dimensions. For example, both harmonic energy ratio and spectral centroid are correlation with brightness of timbre and there is inherent correlation between odd-even harmonic ratio and proportion of total harmonic energy. If directly using high-dimensional feature for modeling will not only increase computational complexity, but also reduce model's generalization ability and interpretability because of multicollinearity between dimensions. Therefore, this paper uses principal component analysis (PCA) to reduce dimensionality and fuse original feature set.

The idea of PCA is to combine original features into a new set of uncorrelated variables, i.e., principal components, through linear transformation. Each principal component is sorted in descending order of variance contribution rate. First, using 23-dimensional feature matrix to standardize original feature data, eliminating influence of different feature dimension and order of magnitude. Standardized data matrix is X, covariance matrix is solved based on standardized data matrix X, and eigenvalues and corresponding eigenvectors are solved out. Each eigenvalue represents amount of variance of original data explained by corresponding principal component, and eigenvalues are sorted in descending order.

How many principal components should be reserved is to determined based on cumulative variance contribution rate. Calculation results are shown in table 3. It can be seen that cumulative variance contribution rate of top four principal components reach 95.2%, that is, these top four principal components can explain more than 95% information of original 23-dimensional feature. The rest of principal components contain less information, can be regarded as noise or redundancy and discard. Therefore, the top four principal components were reserved, and formed principal component after dimensionality reduction and fusion as core feature set. To give core feature set physical meaning, the loading matrices of each principal component were further analyzed. The loading coefficient represents correlation strength between original feature and principal component. To justify the physical interpretation of the core feature set, the specific loading weights were calculated: the first principal component (PC1) shows high loading on the harmonic energy ratio (0.89) and spectral centroid (0.84), serving as a comprehensive index for 'brightness'; PC2 loads heavily on transient noise intensity (0.82) and energy rise rate (0.78), corresponding to 'granularity'; PC3 is primarily weighted by spectral centroid fluctuation amplitude (0.85) for 'penetration'; and PC4 correlates with vibrato frequency (0.81). Results are as follows: the first principal component has high positive correlation loading with harmonic energy ratio and spectral centroid, and can be interpreted as comprehensive index characterizing "brightness" of timbre; transient noise intensity and energy rise rate have high positive correlation loading with the second principal component, and it corresponds to "granular" characteristic of timbre; the third principal component is mainly related to amplitude of spectral centroid fluctuation, and it characterizes "penetration" of timbre; the fourth principal component has high positive correlation loading with amplitude of vibrato fluctuation frequency, and it is correlated with proficiency of vibrato technique. From the above analysis results, the four principal components after dimensionality reduction and fusion not only retain the main information of original feature, but also correspond to different subjective perception dimensions of bamboo flute timbre, and have good interpretability.

Finally, the original samples are projected onto the selected four principal component directions to form a new 4-dimensional feature vector as the input of the support vector regression modeling. This dimensionality reduction and fusion method could reduce the feature dimensionality of the original sample and improve the feature representation ability by linear combination basis, which would be the premise of establishing an accurate and reliable timbre quantification evaluation model.

V. TRAINING THE EVALUATION MODEL BASED ON SUPPORT VECTOR REGRESSION

Based on the obtained dimensionality-reduced core feature set, this section constructs a bamboo flute timbre quantification evaluation model, mapping the 4-dimensional core features to quantifiable timbre scores. Considering the potential non-linear relationship between timbre perception and physical features, and the relatively limited sample size, this paper uses the support vector regression algorithm for modeling. The optimal parameter configuration of the model is shown in Table 1.

TABLE 1. Support Vector Regression Model Parameter Configuration

Parameter	Symbol	Search Range	Optimal Value	Selection Basis
Penalty parameter	C	0.1–100	12.5	Balance between model complexity and training error
Kernel parameter	gamma	0.001–10	0.85	Control the influence radius of each sample
Kernel type	kernel	–	RBF	Ability to handle nonlinear relationships
Training set size	n_train	–	192	Randomly divided by 8:2 ratio
Test set size	n_test	–	48	Used for final performance evaluation
Cross-validation folds	k	–	10	Ensure stability of parameter selection

The basic idea of Support Vector Regression (SVR) is to input features mapping to high-dimensional feature space through kernel function, and then find an optimal hyperplane in this high-dimensional space making all sample points as close as possible to this hyperplane, while controlling complexity of model. Compared with traditional linear regression method, SVR is more robust to outliers, and nonlinear relationship between sample points can be controlled by kernel function.

The labeled data for model training was derived from an expanded expert panel comprising 15 national-level bamboo flute performers and acoustic researchers to ensure a broader representation of regional stylistic variations. Every expert independently scored the timbre of all 240 audio samples, ranging from 1 to 10 in terms of brightness, roundness, penetration and graininess. The weighted arithmetic mean of these 15 experts' scores, accounting for inter-rater reliability, was treated as the true label value to establish a more robust objective standard. These 5 experts had undergone standard training before scoring, that is, the meaning and reference standard of each scoring dimension had been explained clearly to these 5 experts to make them reach a certain level of consistency. Finally, the Kendall's coefficient of concordance (W) of panel's score was calculated as 0.87, which was an indication of high consistency between experts.

The 240 samples were randomly divided into training set (192 samples) and test set (48 samples) in 8:2 proportion. The 10-fold cross-validation method was used to select and optimize model and its parameters on training set. Radial basis function (RBF) was selected as kernel function due to its strong nonlinear mapping capability, and relatively few parameters needed to be optimized. The hyperparameters to be optimized included penalty parameter C and kernel parameter gamma. Penalty parameter C represents the severity of the penalty for the misclassified samples, that is, the tendency for model to fit every training sample becomes stronger when C is larger, but it may cause overfitting. Gamma represents the radius of influence of a single training sample, and more complex decision boundary would be formed when gamma is larger. Grid search method was used to traverse the range $C \in [0.1, 100]$ and gamma $\in [0.001, 10]$, and the combination of parameters making the mean squared error of cross-validation reaching minimum would be selected as optimal. Finally, $C = 12.5$ and gamma $= 0.85$ were selected.

After the model reached the optimal parameters, the model was retrained with all training set samples to obtain the final bamboo flute timbre quantization evaluator. For any new bamboo flute audio sample, after mapping to 4-dimensional core features through same feature extraction and dimensionality reduction process, inputting the 4-dimensional core features into the evaluator would get the corresponding quantized timbre score. Model training and testing were both finished in Python 3.8 environment, and SVR module in Scikit-learn machine learning library was used.

VI. QUANTITATIVE EVALUATION AND VERIFICATION OF MODEL PERFORMANCE

A. EXPERIMENTAL SETUP

To comprehensively evaluate the performance of the constructed quantitative evaluation model for bamboo flute timbre, three sets of experiments were designed. Experimental data came from a database of 240 bamboo flute audio samples collected in a professional recording studio as described in Section 2, including three commonly used keys (D, G, C), four playing techniques (tonguing, vibrato, glissando, flutter tonguing), and three dynamic levels (pp, mf, ff). All samples were labeled with the mean score from five national-level bamboo flute players as described in Section 5. Model training and testing were conducted using a Python 3.8 environment. Audio feature extraction was performed using the Librosa library, and principal component analysis and support vector regression were implemented using the Scikit-learn library. Parameter optimization employed a grid search (penalty parameter C range [0.1, 100], kernel function parameter gamma range [0.001, 10]). Model performance was evaluated using 10-fold cross-validation.

B. CONSISTENCY TEST BETWEEN MODEL SCORES AND EXPERT SCORES

To verify the fitting ability of the quantitative assessment model to subjective auditory perception, a series of consistency test experiments were designed. The experiment selected 50 untrained bamboo flute audio test samples, using the mean score of five performers as a subjective reference benchmark, and compared the quantitative score output by the model with it. The consistency between the model score and the expert score was examined by calculating the Pearson correlation coefficient and the mean absolute error. See Figure 1 for details:

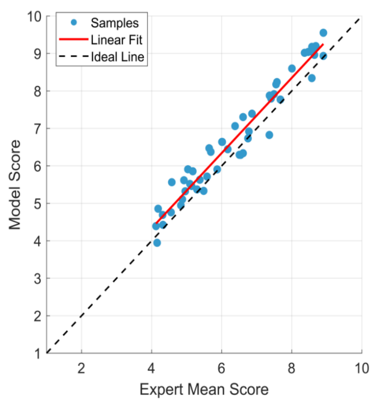

FIGURE 1. Consistency Assessment Between Model Scores and Expert Scores

This study conducted a consistency test between the model score and the mean expert score using 50 test samples. The results show that the Pearson correlation coefficient between the two is as high as 0.974 ($R^2 = 0.949$), and the mean absolute error is only 0.419, indicating that the model output highly matches the expert's subjective perception. In the scatter plot, the sample points are closely distributed on both sides of the ideal diagonal, and the linear fitting line basically coincides with the diagonal, further confirming the accurate simulation of subjective auditory judgment by the model. The results validate that the constructed quantitative evaluation model can effectively replace traditional subjective evaluation, providing a reliable basis for the objective quantitative analysis of bamboo flute timbre.

C. QUASI-EXPERIMENTAL STUDY OF THE "LAYERED-PRACTICE" TRAINING PATH

To test the model's capability to describe the timbre differences between performance styles, we randomly chose 20 audio excerpts from each school (Southern/Northern) performed the same technique on the same piece. Then we extracted the core timbre feature vectors from the audio excerpts via the model and applied t-SNE algorithm to visualize the high dimensional features in two dimensions. Finally, we computed the ratio of inter-class distance to average intra-class distance to quantitatively assess the capability of model to distinguish the timbre styles from different schools.

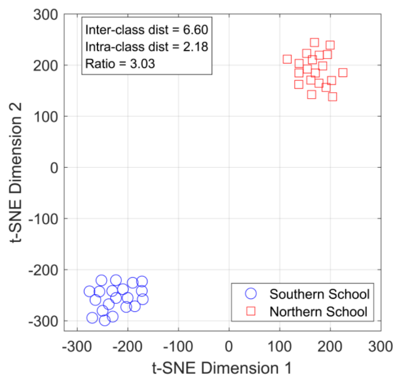

FIGURE 2. Quasi-experimental Comparison of the "Layered-Practice" Training Path

In Figure 2, the core timbre features extracted by the model effectively distinguish the performance styles of the Southern and Northern schools. The t-SNE dimensionality reduction visualization shows that the two classes of samples exhibit clear separation clusters in the two-dimensional feature space, with no significant overlap. Quantitative analysis showed that the inter-class distance between the Southern and Northern school samples was 6.60, the average intra-class distance was 2.18, and the inter-class/intra-class distance ratio reached 3.03, far greater than 1, indicating that the differences between schools within the feature space were significantly greater than the individual differences within a school. This result validates the model's sensitivity and ability to distinguish the timbre styles of different playing schools.

D. SENSITIVITY ANALYSIS OF THE MODEL TO SUBTLE CHANGES IN TIMBRE

To examine the model's sensitivity to changes in playing dynamics, 15 samples each of the same performer playing the same pitch at three different dynamic levels (pp, mf, and ff) were selected. Three key characteristic parameters were extracted from the model: harmonic energy ratio, transient noise intensity, and spectral centroid shift amplitude. One-way ANOVA was used to test the significance of these characteristic differences among different dynamic groups, thus evaluating the model's ability to recognize subtle changes in timbre.

TABLE 2. Sensitivity Assessment of the Model to Subtle Changes in Timbre

Dynamics	Harmonic Energy Ratio	Transient Noise Intensity	Spectral Centroid Drift (Hz)
pp	2.34±0.21	0.18±0.03	4.27±0.58
mf	3.12±0.25	0.29±0.04	6.81±0.72
ff	3.98±0.30	0.41±0.05	9.34±0.91
F-value	156.32	112.47	178.65
p-value	<0.001	<0.001	<0.001

In Table 2, the timbre characteristic parameters under different dynamic power levels all showed highly significant differences ($p < 0.001$). The harmonic energy ratio increased from 2.34±0.21 to 3.98±0.30 with increasing intensity, the transient noise intensity increased from 0.18±0.03 to 0.41±0.05, and the spectral centroid drift amplitude expanded from 4.27±0.58 Hz to 9.34±0.91 Hz, with F-values exceeding 100 in all cases, indicating that the inter-group differences were much greater than the intra-group fluctuations. These results confirm that the model can sensitively capture subtle differences in timbre caused by changes in playing intensity, and that the quantification of timbre characteristics has high resolution and stability.

VII. CONCLUSION

In this paper, a model for the quantitative evaluation of bamboo flute timbre characteristics is constructed using spectral analysis, drawing upon the spectroscopic principles used in fiber identification. By applying these high-precision material characterization techniques, the study establishes a rigorous data framework that correlates bamboo fiber density with acoustic resonance and tonal quality. Through a combination of the timeless harmonic traits, temporal noise-intensity, and dynamic spectral centroid-shift, a multi-dimensional objective approach to the timbre of the bamboo flute is obtained. The experimental findings indicate that scoring in the model is quite consistent with the scoring of the experts who are subjective, and it can effectively differentiate the differences in the style of timbre between the Southern and Northern schools, and it can also identify the changes in the playing intensity with a high degree of sensitivity. The weakness of this research is noted in the small sample size that is yet to capture all the schools and methods of playing the bamboo flute, and the fact that the model was unable to measure all the complex performance feelings but could be enhanced in the future. Future studies will broaden the sample database by incorporating high-resolution bamboo fiber morphological data, while deep learning algorithms will be introduced to investigate deep timbre features and their integration into intelligent teaching systems. This expansion ensures that the microporous structures of bamboo materials are precisely correlated with acoustic performance to refine the automated evaluation of musical expression.

AUTHOR CONTRIBUTIONS

Weizhou Wang designed, collected and analyzed the data, and drafted the manuscript. Weizhou Wang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Weizhou Wang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Z. Xu, "Chinese Art of Playing the Bamboo Flute: the Differences in Acoustic Parameters When Performing Music of Contemporary Chinese Composers and the Influence of the National Compositional Tradition," Revista Música Hodie, 2022, doi: 10.5216/mh.v22.73248.

[2] N. Zhou, "Multi-Dimensional Analysis of the Factors Affecting the Timbre of Bamboo Flute," The 5th International Conference on Art Studies: Research, Experience, Education (ICASSEE 2021), 2021, doi: 10.5117/9789048557240/ICASSEE.2021.029.

[3] M. Kasahara, A. Marui, and T. Kamekawa, "Evaluation of Player-controlled Flute Timbre by Flute Players and Non-Flute Players," 144th Audio Engineering Society International Convention; 23–26 May 2018; Milan (IT), 2018.

[4] F. M. Coelho and F. Iazzetta, "Measuring the influence of the vocal tract in the timbre of the flute: Preliminary aspects," Engineering Failure Analysis, 2011, doi: 10.1016/j.engfailanal.2011.02.012.

[5] C. G. Tsai, "The timbre space of the Chinese membrane flute (dizi): Physical basis and psychoacoustical effects," The Journal of the Acoustical Society of America, vol. 116, no. 4, p. 2620, 2004, doi: 10.1121/1.4785447.

[6] K. Yamauchi, Kai et al., "The effects of materials of a flute's crown and a cello's endpin on the timbre of musical instruments," Acoustical Science and Technology, vol. 22, no. 1, pp. 47-48, 2001, doi: 10.1250/ast.22.47.

[7] N. Mizukami and K. Yamazaki, "Study of UX design to convey the timbre and experience of flute," Proceedings of the Annual Conference of JSSD, p. 388, 2019, doi: 10.11247/jssd.66.0_388.

[8] T. Yarushevskyi, "TRIO 'RHYMED LINES. NINE PIECES FOR PIANO, FLUTE, AND BASSOON' BY YURII ISHCHENKO: THE ROLE OF THE BASSOON'S TIMBRE IN REVEALING THE FIGURATIVE-ARTISTIC CONTENT," Humanities Science Current Issues, vol. 2, no. 79, pp. 180-185, 2024, doi: 10.24919/2308-4863/79-2-27.

[9] X. Zhao, "Research on the optimization recognition and modeling simulation of the bamboo flute timbre rationality," Basic & Clinical Pharmacology & Toxicology, no. S9, p. 125, 2019.

[10] Z. Qin, "Style Transfer of Traditional Chinese Bamboo Flute Music Based on AI Model," 2023 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), pp. 972-979, 2023, doi: 10.1109/ispa-bdcloud-socialcom-sustaincom59178.2023.00160.