Sound Analysis of Violin via Neural Network and Numerical Simulation

Masao Yokoyama
Meisei University
masao.yokoyama@meisei-u.ac.jp

Yuki Hori
Bunkyo-Gakki Co.
hori@bunkyo-gakki.com

Shouta Tani
Miyazaki University
hl19051@student.miyazaki-u.ac.jp

Amane Takei
Miyazaki University
takei@cc.miyazaki-u.ac.jp

Genki Yagawa
Toyo University, University of Tokyo
yagawag@gmail.com

ABSTRACT

Two approaches, a neural network and numerical simulation, for analyzing the sound of violins are introduced herein. First, the sounds (open strings and music) of more than 20 violins are recorded. A set of acoustic features, such as the spectrum envelope and mel-frequency cepstrum coefficients, is used to train the neural network and identify the violin timbre. The possibility of quantifying the similarity of the violin timbre and identifying the violin maker is demonstrated. Subsequently, the violins are scanned using a micro-computed tomography scanner to retrieve geometric data, and numerical simulations of the vibration and sound radiation of the violin body are performed. The vibration of the top plate due to the sinusoidal oscillation of the bridge and the sound radiation from the f-hole and C-bouts are calculated using finite-element software COMSOL Multiphysics and ADVENTURE-Sound.

1. INTRODUCTION

Recently, numerical simulations using the finite element method (FEM) have been performed to analyze violin vibrations and sound radiation [1, 2]. To obtain the three-dimensional geometric data of the violin shape, a laser scanner [3] and computed tomography (CT) scanner [4] were used. As shown in these studies, the analysis of violin sounds using computers is flourishing. Numerical simulation is effective for qualitatively and quantitatively modeling vibration and acoustic features. Meanwhile, the timbre of violins is an interesting topic and has been investigated extensively. For example, Setragno [5] and Fritz [6] compared old and modern violins, whereas Corradi et al. [7] and Schleske et al. [8] conducted multidisciplinary analyses of violins.

We conducted a numerical simulation of old violins used by masters, such as the Stradivarius, using the FEM, and modeled the effects of wood properties on the vibration modes of the violin body [9]. Geometric data for the violins were obtained using a micro-CT scanner. Analysis using numerical simulations and CT scanners is a noninvasive and nondestructive method for analyzing historical assets such as Stradivarius violins. Furthermore, we can analyze the sound features of violin using the numerical simulation quantitatively and qualitatively without making lots of violins.

In this study, we recorded the sounds of more than 20 violins, from antiques to brand-new violins. We are currently developing a system that can identify the timbres of violins using neural networks. We attempted to determine whether a computer can identify differences in the timbre of violins as violinists and dealers do. The possibility of identification will contribute to the appraisal and authorization of masters' violins and improve the sound quality of electric instruments and MIDI.

Herein, we introduce our approach for analyzing the sound features of a violin using neural networks and numerical simulations.

Copyright: © 2023 Masao Yokoyama et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2. RECORDING

A list of the violins investigated in this study is shown in Table 1, which includes old Italian violins from the 17th century to the brand-new Japanese commercial violins. We recorded the sound played by a violinist on open strings without any musical expression (E5, A4, D4, G3, long-tone of approximately 4 s), the G major scale, and the music piece "Meditation from Thaïs" with musical expression (including vibrato, dynamics, diminuendo, and crescendo).

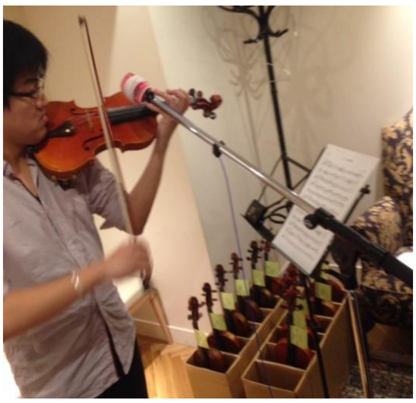


Fig. 1 Recording violin sound using fast Fourier transform analyzer in a rehearsal room.

The piece was played at a slow tempo and had several long notes, and the pitch range was wide, i.e., from A3 (220 Hz) to F#6 (1480 Hz).

A fast Fourier transform analyzer (Oros NV Gate) and a nondirectional condenser microphone (ICP 1/4-inch microphone) were used for recording (Fig. 1). A microphone was placed approximately 10 cm above the violin bridge. Each violin was played twice in a small rehearsal room with reduced echo. The strings of all violins were of the same brand (dominant (A, D, G) and Goldbrokat). Additionally, the same bow and rosin were used.

Violin maker (country)	Year
Catenali (Italy)	ca. 1690
Stradivari (Italy)	1698
Pietro Guarneri (Italy)	ca. 1700
Santo Serafin (Italy)	ca. 1700
Gragnani (Italy)	1760
Balestrieri (Italy)	1780
Pressenda (Italy)	1838
Fabris (Italy)	1870
Scarampella (Italy)	1907
Fagnola A(Italy)	1923
Fagnola B(Italy)	1931
Genovese (Italy)	1927
Michetti (Italy)	1929
Guerra (Italy)	1941
Bisiacchi (Italy)	1953
Garinberti (Italy)	1967
Contemporary violin middle-class (Japan)	2015
Contemporary violin Economic (Japan)	2015
Contemporary violin Stradivari Copy (Japan)	2015
Contemporary violin Guarneri Del Gesu Copy A (Japan)	2015
Contemporary violin Guarneri Del Gesu Copy B (Japan)	2015

Table 1. List of violins used for recording.

3. TIMBRE IDENTIFICATION VIA NEURAL NETWORK

3.1 Identification of timbre based on spectral envelope

The spectral envelope is one of the most important pieces of acoustic information in speech-signal processing. The envelope curve of the power spectrum and local peaks were observed by calculating the power spectrum of the violin sound wave. The peaks emerging in the envelope curve of the spectrum were named the first, second, third, and fourth peaks from lower to higher frequencies (labeled F1–F4, respectively, in this study). Woodhouse [10] mentioned that peaks (humps) over 1 kHz, which are known as transition peaks and bride hill, are similar to vowel formants. For example, the formant frequencies of human vowels “a”’ and “e” are different, and we can distinguish an “a” from an “e.”’ Therefore, we applied this property to the identification of violin timbres. We expect that the differences in the timbre can be explained by the patterns of these peak frequencies.

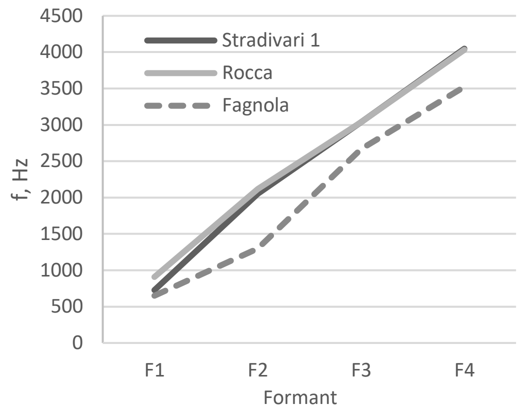


Fig. 2 Distribution patterns of peak frequencies (F1–F4) of power spectrum in open string A[11]. The patterns of the Stradivari and Rocca are similar, whereas that of the Fagnola is different from those of the Stradivari and Rocca.

We categorized the peak frequencies (F1–F4) of the recorded sound, as shown in Fig. 2, using the Praat (a sound analysis software) [11]. Some patterns were observed, including linear straight, zigzag, and downward curving patterns. Fig. 2 shows an example of the analysis results, where the patterns of the Stradivari and Rocca violins were almost straight and close to each other, whereas the pattern of the Fagnola violin was zigzagged.

Furthermore, we conducted a perceptual test using a questionnaire to distinguish the timbres of the recorded violin performance [11]. In the test, the participants (10 professional violinists/dealers) listened to three sound waves. After listening to two (A, B) of the three sounds, the participants listened to the third sound (X) and answered whether the timbre of X resembled A or B.

Based on the experimental results, in the case of the sounds of two violins with different distribution patterns, such as straight and zigzag patterns, the participants perceived that the two violins were not similar in terms of timbre. However, when the distribution patterns of the two violins were similar and the difference was insignificant, the participants answered that the two violins were more similar. Thus, the experiment revealed that the shape of the distribution pattern and the difference between distribution patterns can facilitate the identification of violin timbre.

Using the acoustic characteristics of violin sounds, we attempted to determine whether a computer can identify the differences in the timbre of violins. By learning significant amounts of sound data, one may be able to identify certain violins using machine learning techniques.

The acoustic features, i.e., the spectrum envelope and mel-frequency cepstrum coefficients (MFCCs), of the recorded sound data were calculated using the Python software. Neural network training was performed using a pair of acoustic features and violin markers, where the input was the acoustic features and the output was the violin marker. Finally, the trained network was applied to the test data to validate its accuracy and to calculate the similarity for each violin. An identification label was assigned to each of the 21 violins shown in Table 1 and to four strings; thus, 84 labels were assigned to distinguish the timbres.

The neural network was a fully connected four-layer network. The number of inputs was 1024, and in the case where the training data was a spectrum envelope, the number of neurons in the second and third layers was 512. The number of outputs in the fourth layer was the same as that of the labels. The program for calculating the cepstrum as preprocessing and to execute machine learning was written using Python and the Keras library, which is the front end of TensorFlow. The activation function was the ReLU function. The learning rate was 0.1 and the dropout ratio was 0.2. The loss function was a categorical cross-entropy.

3.2 Comparison of accuracies

Fig. 3 shows a comparison of the accuracies of the acoustic features (spectral envelope, mel spectrum, and MFCC). As shown, MFCC is the best option for identifying violin sounds. Additionally, the accuracy was high (approximately 90%) when the neural network was trained and tested using only open-string data. This is because the sound wave of an open string is approximately periodic. However, when a musical piece with musical expression was used as the training data, the accuracy decreased (see middle of the graph in Fig. 3, without string distinction) because the musical piece contained a complex set of parameters, including vibrato and dynamic changes.

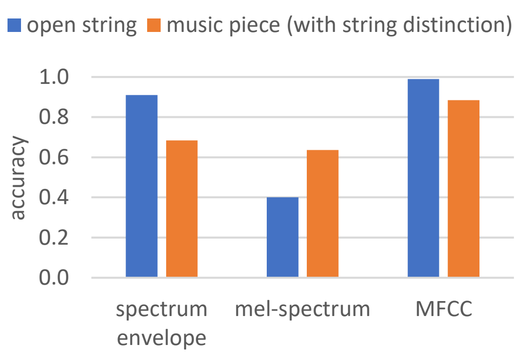


Fig. 3 Accuracy based on nature of sound.

3.3 Similarity in timbre

The probabilities predicted by the proposed neural network are shown in Fig.4. The MFCC data for the four violins selected from Table 1, including Stradivari and Fagnola (shown in legend under the graph) were used to train the neural network. Violins represented on the vertical axis were used to predict the similarity and were assumed to be unknown.

For example, our program predicted that Michetti would be similar to Stradivari by 45%, whereas Gragnani would be similar to P. Guarneri by 83%. However, the timbres of the two band-new violins, the Master-new Del Gesu copies A and B, resembled each other. These two violins are copies of the Guarneri Del Gesu, which was made by the same Japanese master. In addition, Fagnola B was similar to Fagnola A. Interestingly, the sounds of violins from the same maker using the same copy model resembled each other.

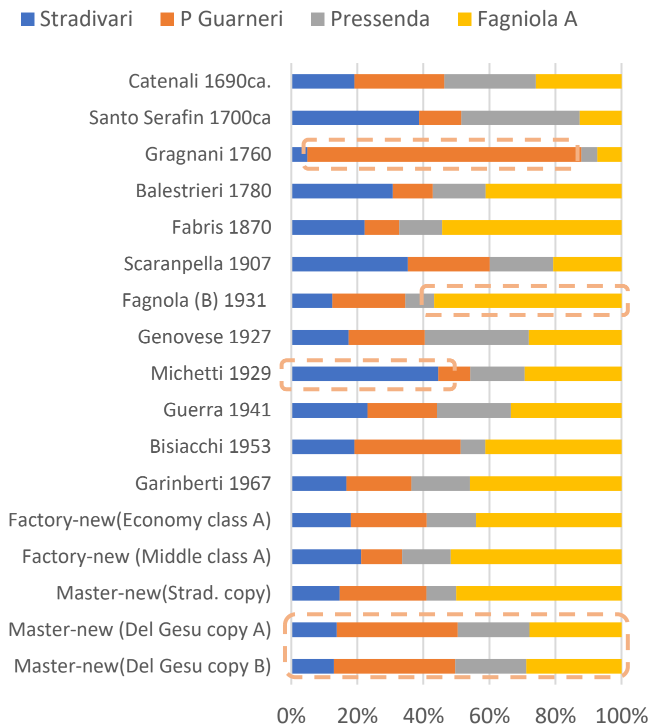


Fig. 4 Percentages of the timbre similarity of test violins against eight famous violins (Dataset: Performance of a music piece, MFCC). Michetti is similar to Stradivari by approximately 45 %. Fagnola B is similar to Fagnola A.

Although this result is only an example, we expect that a neural network can indicate the similarity among violins quantitatively and distinguish the timbre of a violin by learning acoustic features, such as the MFCC and spectrum envelope. As an application of this system, when violinists purchase a violin at a shop, the system may provide a useful visualization for quantitatively determining sound features. Violinists who wish to purchase a violin that produces sounds similar to the Stradivari can select one that is the most similar to Stradivari in terms of percentage. In addition, we may be able to develop an appraisal machine using sound to authenticate Stradivari violations in the future.

However, we used only one violin from each violin manufacturer in this experiment. Hence, we cannot definitively conclude that the trained model of the network can express the general characteristics of a violin manufacturer's timbre. More recorded data per violin manufacturer are required for a correct appraisal.

4. SCANNING OF GEOMETRIC DATA USING MICRO-CT SCANNER

The geometries of violins made by Stradivari and Guarneri del Gesu were scanned using a micro-CT scanner, whose precision was 0.1 mm. The upper photograph in Fig. 5 shows a cross-sectional image of the violin body obtained via a micro-CT scan. The lower image in Fig. 5 shows an image of the interior of the violin body (Stradivari, 1719) using scanned geometric data and computer-assisted design (CAD) software.

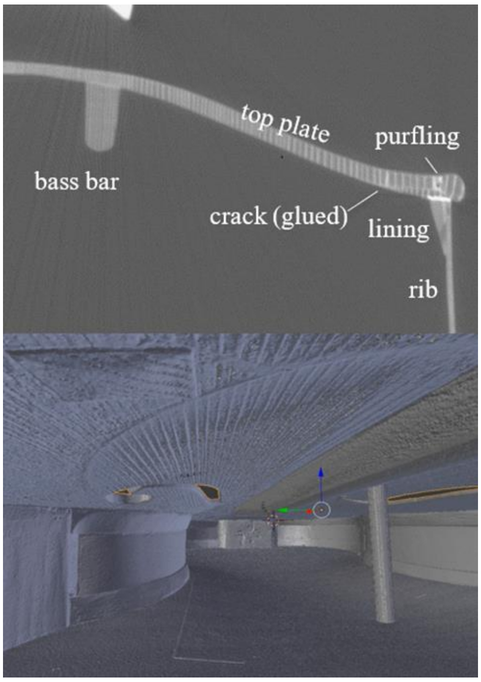


Fig. 5 Visualization of interior of violin via micro-CT scanning (upper side) and CAD software (lower side).

Because the scanned raw geometric data included many fragments and holes, we cleaned the data using the CAD software before performing numerical simulation and visualization. Using this scanner allowed us to observe not only the details inside the violin body, but also the grains of the wood, cracks, and traces of restoration. The geometric data were classified into components such as top and back plates, ribs, sound posts, and bass bars such that different mechanical properties can be specified for each component. The scanned data were saved as STEP files (the standard for exchanging product model data).

5. NUMERICAL SIMULATION OF SOUND RADIATION OF VIOLIN

The import of geometric data, meshing, and FEM calculations were conducted using COMSOL Multiphysics^{TM}[12]. The STEP files created (as mentioned in Section 4) were imported into the COMSOL Multiphysics software as geometric objects (Fig. 6, upper). In addition, a spherical area of air surrounding the violin was constructed (Fig. 6, lower).

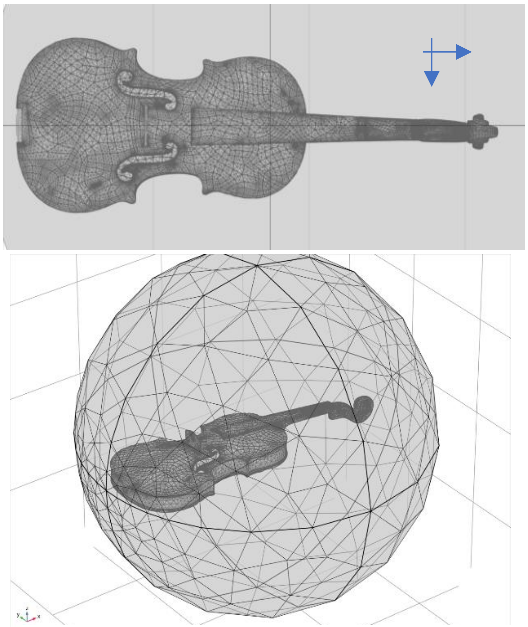


Fig. 6 Mesh of violin and air field generated using auto-mesh function in COMSOL Multiphys-

In COMSOL Multiphysics, the mesh generator discretizes the domains into tetrahedral second-order mesh elements using the free-mesh method. Approximately two million elements were generated, including those for the violin and air. The eigenfrequency, body displacement, and sound pressure were calculated using the FEM via the acoustic–structure interaction module in COMSOL Multiphysics.

The mechanical characteristics of the wood for the violin were set in three orthogonal directions in COMSOL Multiphysics (longitudinal grain direction, radial annual ring direction, and direction tangential to the annual ring). We specified the values for the mechanical properties, such as the Young's modulus, rigidity modulus, and Poisson's ratio, based on measurement values by Green et al. [13]. We set the representative values of density for maple and spruce as 0.63 and 0.36, respectively, and their Young's moduli as 12.6 and 9.9 GPa, respectively.

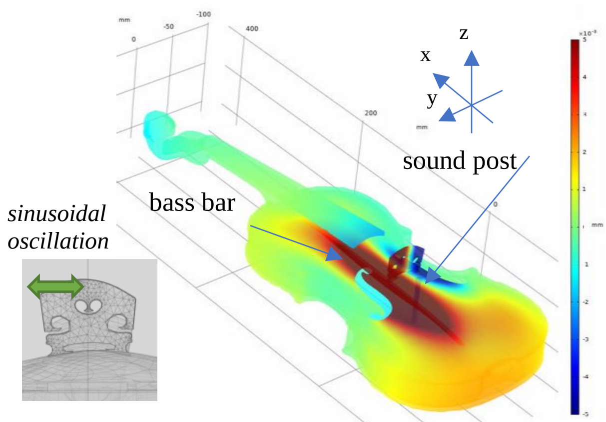


Fig. 7 Displacement of violin by forced vibrations on the bridge, where the G string is placed (196 Hz). Bass-bar side shows significant vibrations. Scroll and fingerboard vibrate as well.

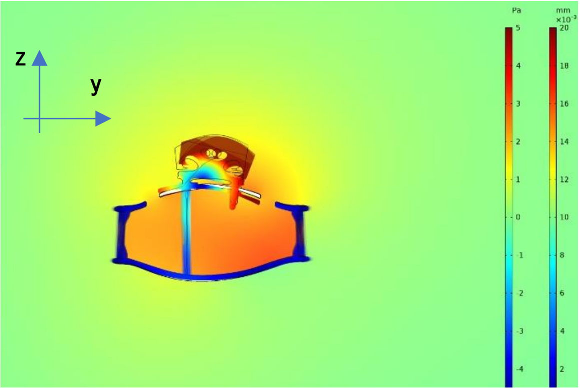


Fig. 8 Visualization of displacement of bridge and sound pressure field in the y–z plane.

5.1 Vibration of violin body

Figure 7 shows the simulation results when a forced sinusoidal oscillation was applied to the bridge of the violin. The color contours depict the displacement of the body in the z-axis direction caused by forced vibrations. The left side of Fig. 7 shows the oscillating position of string G. The frequency of the sinusoidal function at 196 Hz (G3, the fundamental frequency of the G string) was input along the y-axis.

The bridge alternately oscillated from side to side along the y-axis, and alternate vibrations were induced on the top plate by the bridge oscillation. In particular, the magnitude of the displacement on the bass-bar side (left side of the violin) was significant. We speculated that the vibrations of the scroll and fingerboard did not significantly affect the sound volume and timbre; however, these vibrations were not negligible, as shown in the video depicting the detailed simulation results [14].

5.2 Near sound field

The acoustic pressure fields in the y–z plane (including the bridge and sound post) and the x–y plane (30 mm above the arch of the top plate) are shown in Figs. 8 and 9, respectively. Based on the vibration of the top plate caused by the sinusoidal oscillation of the bridge, we simulated the sound radiation in a concentric circle from the f-hole and C-bout. The result obtained was similar to the experimental result obtained by Wang [15], in which sounds at low pitches radiated concentrically from the violin body. In addition, as shown in Fig. 8, the area of the bass bar fluctuated when the sound post was a fulcrum. The sound post functioned as a fulcrum of vibration and a conduit for connecting the vibration between the top plate and backplate.

5.3 Expansion for radiation in a hall using parallel computer

In using the FEM to calculate the sound radiation in a wide area, such as a chamber room or concert hall, billions of meshes are required. Such calculations are difficult to perform on a personal computer; therefore, a large-scale parallel computer is necessitated. We plan to expand our numerical simulation to calculate sound radiation in a concert hall using the parallel computer software, ADVENTURE Sound [16].

The acoustic pressure and coordinate data on the violin surface obtained using COMSOL were transferred to the ADVENTURE Sound software, which is a parallel acoustic analysis code. An iterative domain decomposition method was applied to ADVENTURE Sound. IDDM is the most efficient parallel technique for large-scale analysis, which features several hundred million to several billion degrees of freedom implemented using the hierarchical domain decomposition method [17].

The violin was constrained in a spherical air domain, and the analysis results obtained using ADVENTURE Sound based on a frequency of 196 Hz at the bridge is shown in Fig. 10. The mesh size was 2.76 mm and the number of meshes was approximately 20 million. As a

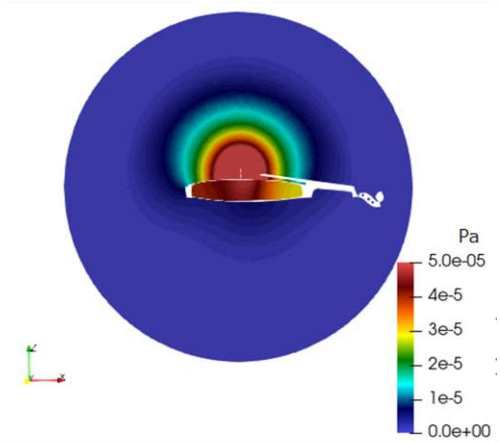
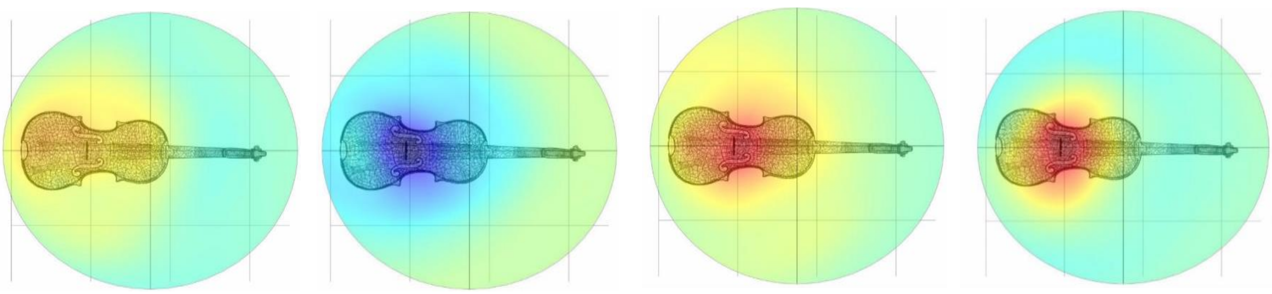


Fig. 10 Visualization of sound radiation around a violin by ADVENTURE Sound.



$t = 0.1T$ ($T = 1/196$ s) $t = 0.15T$ $t = 0.3T$ $t = 0.6T$

Fig. 9 Temporal change in acoustic pressure around violin. T represents a cycle of sinusoidal oscillation on the bridge.

boundary condition, sound absorption was set at the surface of the sphere. A PC cluster (Intel Corei7-9700K, 3.60 GHz, 32 GB RAM/Node, nine nodes) was used for the calculations. The results indicate that the radiation of sounds from the violin surface to the air domain around the violin can be calculated.

Currently, a program for large-scale analysis using acoustic pressure on the surface of a violin body is being implemented. Dynamic analysis of sound radiation from forced vibrations on a violin bridge in a hall will be conducted in the near future.

6. CONCLUSIONS

In this study, experiments using neural networks and numerical simulations were performed to investigate the timbre of violins. We demonstrated the possibility of predicting violin manufacturers based on the acoustic features (spectrum envelope and MFCC) of violins using a deep learning program. To authenticate Stradivari's violin, the sounds of many violins must be recorded; however, an appraisal machine may be realizable using artificial intelligence.

In addition, we scanned old Italian violins using a micro-CT scanner and performed a numerical simulation of sound radiation from the violin body. Experimentally, the sound pressure around an instrument can be analyzed in an anechoic chamber using array microphones. However, this method is expensive, and securing the appropriate facilities for array microphones is difficult. Therefore, a coupled numerical analysis of the vibrations and acoustics can be substituted for the experiments. We are planning to connect these experimental analysis (using recording and neural network) and numerical analysis as future work.

Acknowledgments

This study was supported by JSPS KAKENHI (grant number 21K11969).

7. REFERENCES

[1] C. Gough, "Acoustic characterization of string instruments by internal cavity measurements" in J. Acoustical Society of America, 150(3), 1922-1933, 2021.

[2] J. A. Torres, C.A. Soto, and D. Torres-Torres "Exploring design variations of the Titian Stradivari violin using a finite element model" in J. Acoustical Society of America, 148(3), 1496-1506, 2020.

[3] M. Pezzoli, R.R. De Lucia, F. Antonacci, and A. Sarti, "Predictive simulation of mechanical behavior from 3D laser scans of violin plates" in Proceedings of International Conference on Acoustics 2019, 2019.

[4] G. Bissinger, and D. Oliver, "3-D laser vibrometry on legendary old Italian violins" in Sound and Vibration 41.7, 10-15, 2007.

[5] F. Setragno, M. Zanoni, F. Antonacci and A. Sarti,. "Feature-Based Timbral Characterization of Historical and Modern Violins" in International Symposium on Musical Acoustics, 90-93, 2017.

[6] C. Fritz et al. "Soloist evaluations of six Old Italian and six new violins" in Proceedings of the National Academy of Sciences 111.20, 7224-7229, 2014.

[7] R. Corradi et al. "A multidisciplinary approach to the characteri-zation of bowed string instruments: the musical acoustics lab in Cremona" in Proceedings of the 22nd International Congress on Sound and Vibration(ICSV22), 2015.

[8] M. Schleske, "Empirical tools in contemporary violin making: Part I. Analysis of design, materials, varnish, and normal modes" in Catgut Acoust. Soc. J., 4(5) 50–64, 2002.

[9] M. Yokoyama, "Coupled numerical simulations of the structure and acoustics of a violin body" in J. Acoustical Society of America, 150 (3), 2058–2064, 2019.

[10] J. Woodhouse, "The acoustics of the violin: a review." Reports on Progress in Physics 77.11, 115901, 2014.

[11] M. Yokoyama, "Possibility of distinction of violin timbre by spectral envelope," in Applied Acoustics, 157, 107006, 2020.

[12] Acoustic-Structure Interaction, COMSOL Homepage, https://www.comsol.jp/model/acoustic-structure-interaction-417.

[13] D. W. Green, J. E. Winandy, and E. E. Kretschmann, "Mechanical properties of wood" in Wood handbook: Wood as an engineering material. Madison, WI: USDA Forest Service, Forest Products Laboratory. General technical report FPL; GTR-113: 4.1-4.45, 113, 1999.

[14] https://youtu.be/ m3cgG-TJs9Q

[15] L.M. Wang and C. B. Burroughs, "Acoustic radiation from bowed violins" in Journal of the Acoustical Society of America. 110(1), 543–555, 2001.

[16] ADVENTURE Project home page: https://adventure.sys.t.u-tokyo.ac.jp/ (Access on 15 Nov. 2021) .

[17] A. Kudo and A. Takei, "Performance evaluation of wave-sound analysis code: ADVENTURE Sound" in Proceedings of the 40th JSST Annual Conference: International Conference on Simulation Technology, 1-3, 2021.