Computational Modelling of Timbre Dimensions for Automatic Violin Tone Quality Assessment

Sergio Giraldo[1†], Rafael Ramirez[1], George Waddell[2] and Aaron Williamon[2]
[1] Music Technology Group, Pompeu Fabra University, Barcelona, Spain
[2] Centre of Performance Science, Royal College of Music, London, United Kingdom
[†] sergio.giraldo@upf.edu

Aims/goals

Automatic assessment of music performance is an open research area widely studied in the past. A vast number of systems aiming to enhance the learning process of a musical instrument are being developed in recent years. However, most of the systems focus on the assessment of pitch and onset accuracy, and very few pay attention to tone quality. This is particularly true in violin music education, where although a consensus exists on what is a good or a bad tone quality, there is not a formal definition due to its subjectivity. We present a machine learning approach for the automatic assessment of violin tone quality. We depart from our previous work on the preliminary modelling of several dimensions involving tone quality. Based on recorded examples of tones with different qualities defined and recorded by a professional violinist, we applied machine learning techniques to learn computational models able to evaluate tone quality from extracted audio features. The tone quality models were implemented into a real-time-visual-feedback system.

Background information

The quality of a performed sound is assumed to be a contribution of several parameters of sound such as pitch, loudness, and timbre. Eerola et al. (2012) identify 26 acoustic parameters of timbre among several instrument groups, which might reflect instrument performance techniques and expressive intentions. Automatic characterization of dynamics and articulation from low-level audio features has been studied by Maestre & Gómez (2005) in the context of expressive music performance. Knight et al. (2011) study the automatic assessment of tone quality in trumpet sounds using machine learning techniques. Romani Picas et al. (2015) make use of machine learning techniques to identify good- and poor-quality notes of the trumpet, clarinet and flute, given training data consisting of low and high level audio features extracted from performed musical sounds with each instrument. Giraldo et al. (2017a, 2017b) conducted perceptual tests on the quality of performed musical notes to study the correlations among previously defined terms used in the literature (i.e., pitch-dynamic-timbre stability, richness, and attack clarity) and a list of tone qualities provided by music experts in violin education.

Tone Qualities	
Dark	Bright
Cold	Warm
Harsh	Sweet
Dry	Resonant
Light	Heavy
Grainy	Pure
Coarse	Smooth
Closed	Open
Restricted	Free
Narrow	Broad

Table 1. Proposed list of tone quality by music experts.

Methodology

We obtained recorded examples of each of the tone qualities listed in Table 1, by a professional violinist. Eight notes were recorded for each tonal label using different combinations of fingerings, strings, and bow directions (i.e. up and down), for a total of 80 notes. Low- and high-level descriptors were extracted from the audio signal using the Essentia library (Bogdanov et al. 2013). Automatic feature selection methods

(i.e., wrapper with genetic search) were used to obtain a subset of low-level frame based, spectral descriptors. After feature selection, machine learning techniques were applied to obtain a model for each tone quality pair in Table 1 (e.g., bright-dark, cold-warm), thus obtaining a total of 10 models. For each model, three machine learning schemes were compared, i.e., Linear Regression (LR), Artificial Neural Networks (ANN) with one hidden layer, and Support Vector Machines (SVM) with linear kernel.

Results

The accuracy of each model was calculated based on the percentage of correctly classified instances. The average accuracy of each of the methods compared, across the 10 models of tonal quality pairs was 79% (5%std), 89% (5%std), 85% (5%std) for LR, ANN, and SVM, respectively. The ANN models were then chosen to be implemented on a real-time system for tone quality assessment, where the probability of each binary class was mapped into a visual feedback system, using a spider chart scheme.

Conclusions

A computational approach to automatically assess the quality of performed violin sounds is proposed. We recorded examples of expert defined tone qualities. High and low-level descriptors were extracted from the audio signal and machine learning models were obtained to predict the different quality dimensions. The accuracy obtained indicate that the selected features contain sufficient information to predict the studied tone qualities.

Acknowledgments

This work has been partly sponsored by the Spanish TIN project TIMUL (TIN2013-48152-C2-2-R), the European Union Horizon 2020 research and innovation programme under grant agreement No. 688269 (TELMI project), and the Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502).

References

Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J. R., Serra, X. et al. (2013), Essentia: An audio analysis library for music information retrieval., in 'ISMIR', pp. 493–498.

Eerola, T., Ferrer, R. & Alluri, V. (2012), 'Timbre and affect dimensions: evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds', *Music Perception: An Interdisciplinary Journal,* 30(1), 49–70.

Knight, T., Upham, F. & Fujinaga, I. (2011), The potential for automatic assessment of trumpet tone quality., in 'ISMIR', pp. 573–578.

Maestre, E. & Gómez, E. (2005), Automatic characterization of dynamics and articulation of expressive monophonic recordings, in 'In *Proceedings of the 118th Audio Engineering Society Convention*, Citeseer.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N. & McAdams, S. (2011), 'The timbre toolbox: Extracting audio descriptors from musical signals', The Journal of the Acoustical Society of America 130(5), 2902–2916.

Romani Picas, O., Parra Rodriguez, H., Dabiri, D., Tokuda, H., Hariya, W., Oishi, K. & Serra, X. (2015), A real-time system for measuring sound goodness in instrumental sounds, in 'Audio Engineering Society Convention 138', Audio Engineering Society.

Giraldo, S., Ramírez, R., Waddell, G. and Williamon, A. (2017a). "A Computational Approach for Measuring Performance Quality in Violin Tones ." In *International Symposium in Performance Science*, 29 August 2017, Reikiavik, Iceland.

Giraldo, S., Ramírez, R., Waddell, G. and Williamon, A. (2017b). "A Real-time Feedback Learning Tool to Visualize Sound Quality in Violin Performances." In *Proceedings of the 10th International Workshop on Machine Learning and Music (MML 2017)*, 6 October 2015, Barcelona, Spain.