Phrase-level modelling of expressive dynamics in violin performances

Fabio Ortega[1], Sergio Giraldo, Rafael Ramirez
Music and Machine Learning Lab, Universitat Pompeu Fabra, Barcelona
[1]Corresponding author: Fabio Ortega, *fabiojose.muneratti@upf.edu*
Submission type: empirical paper

Background. Computational models for expressive music performance attempt to understand and emulate the variations in timing, timbre and dynamics that musicians introduce when performing a musical piece. In the context of music learning, such models provide a valuable tool for teaching students how to play expressively, which is a difficult task from a pedagogical perspective. Very few approaches have been proposed for modelling expressive performance in the violin, and most of them target note-level prediction, missing on the context-sensitive nature of expressive interpretation.

Aims. This study presents a method for automatically estimating plausible expressive variations in dynamics for a violin rendition of arbitrary music scores with the purpose of facilitating expressive performance learning by students. The approach is based on analyzing a database of expressive performances recordings by experts, comparing the performed scores with a new score, and adapting the experts' transformations to render an expressive performance of the new score.

Method. A target musical score is automatically segmented into phrases based on the LBDM algorithm (Cambouropoulos, 2001). For each segment, a search is made in a database of music scores performed by expert violinists to find the closest match according to a melodic distance calculation derived from Stammen and Pennycook (1993). The dynamics curve outline of the match and its relative dynamic level are then applied to the target phrase generating the output. Initial testing has been conducted by recording performances of a violinist for usage as both reference for the predictions and test set for analysis in a leave-one-out approach.

Results. Preliminary results indicate that the predicted dynamics approximate performed dynamics well only when the similarity measurements between desired phrases and matching references remain below a certain threshold. Specifically, dynamics predicted for the segments with close matches show a mean correlation coefficient of 0.5670 ± 0.2243 when compared to dynamics performed by an actual violinist with 95% confidence. Due to the nature of our experimental procedure, there were no attempts to predict the dynamics of one performer based on samples of different ones, nor to observe the effects of a database containing samples of mixed musical genres and recording settings.

Conclusions. The method proposed shows potential for modelling not only dynamics but other aspects of music performance. As the above results suggest, considering a large database of reference performances is very likely to improve the accuracy of predictions. The obtained results seem to corroborate that melodic and rhythmic aspects of musical pieces affect their interpretation by musicians.

Keywords: Music performance, computational modelling, machine learning, violin

References

Cambouropoulos, E. (2001). The Local Boundary Detection Model (LBDM) and its application in the study of expressive timing, Proceedings of the 2001 International Computer Music Conference, Havana.

Stammen, D., & Pennycook, B. (1993). Real-time recognition of melodic fragments using the dynamic timewarp algorithm. Proceedings of the 1993 International Computer Music Conference, San Francisco.