Phrase-level modelling of expressive dynamics in violin performances

*Fabio Ortega1, Sergio Giraldo1, Rafael Ramirez1*

*1*Music and Machine Learning Lab, Universitat Pompeu Fabra, Barcelona

**Background.** Computational models for expressive music performance attempt to predict or emulate the variations in timing, timbre and dynamics that musicians introduce when performing a musical piece. In the context of music learning, such models may provide valuable aid in teaching students how to play expressively, which is a difficult task from a pedagogical perspective. Considering state-of-the-art models available, little work is readily applicable to the violin, and most approaches target a note-level prediction, so the context-sensitive nature of expressive interpretation is not sufficiently represented in these models.

**Aims.** This study presents a method for estimating plausible variations in dynamics for a violin rendition of a new score with the purpose of facilitating expressive performance learning by students. The approach consists basically of comparing melodic and rhythmic content of phrases in the score against a database of expert performances and adapting best matches.

**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 melodies played by expert performers for the closest match according to a musical 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. Especifically, 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, whereas the same measurement for the entire sample (not necessarily having close matches) does not differ from zero significantly. 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, a large database of reference performances should directly impact the quality of predictions, but we have yet to test how large it should be for practical uses. More importantly, successful application of our approach reinforces the belief that melodic and rhythmic aspects of a musical phrase affect its interpretation by musicians.

**Keywords:** Music performance, computational modelling, 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.