

Measuring ensemble synchrony through violin performance parameters: a preliminary progress report

Panagiotis Papiotis, Marco Marchini, Esteban Maestre and Alfonso Perez

Music Technology Group, Universitat Pompeu Fabra
Tanger 122-140, 08018 Barcelona, Spain
{panos.papiotis, marco.marchini, esteban.maestre, alfonso.perez}
@upf.edu

Abstract. In this article we present our ongoing work on expressive performance analysis for violin and string ensembles, in terms of synchronization in intonation, timing, dynamics and articulation. Our current research objectives are outlined, along with an overview for the methods used to achieve them; finally, focusing on the case of intonation synchronization in violin duets, some preliminary results and conclusions based on experimental recordings are discussed.

Keywords: violin, expressive performance, intonation, ensemble performance, bowing gestures, motion capture

1 Introduction

Expressive music performance analysis is a wide and interdisciplinary research field, combining elements from signal processing, computational musicology, pattern recognition, and artificial intelligence among others. A thorough presentation of the state of the art in the field can be found in [1].

The problem of expressive performance analysis can be stated as follows: given a score S_i and a recorded performance of that score E_i , performance analysis is carried out by measuring the deviations between S_i and E_i in terms of *timing* (onset times and note durations), *dynamics*, *articulation* (vibrato, tremolo, overall timbre etc.) and, depending on the instrument with which the piece is performed, *intonation*. These deviations can then be seen as the performer's interpretation of the piece, a combination of personal artistic choices as well as implicit musical knowledge.

In the case of ensemble music, where multiple musicians are performing their respective parts simultaneously, performance analysis can be performed on two different levels: on the *intrapersonal* level, thus studying only the relationship between each musician's individual performance and the score it is based on, and on the *interpersonal* level [2], where the relationship between the musicians' performances is also studied; specifically, the influence of one musician's specific performance parameters (either performed live or pre-recorded) on the specific performance parameters of each other member of the ensemble. Naturally, this case is significantly less studied than the case where the focus is placed on a single musician,

although in the past decade a number of researchers have investigated the subject from various viewpoints [3], and utilizing different performance parameters [4].

1.1 Objectives

In studying the influence of one musician over the other (and vice versa) in an ensemble situation, there are three main objectives to be regarded:

- *Detecting* the source and target of the influence mechanism,
- *Analyzing* the nature of the influence mechanism, and
- *Simulating* the influence mechanism by means of a computational model.

It is important to note that the first of these objectives, namely *detecting* the source and target of the influence mechanism, begins with the hypothesis that such an influence actually exists. Therefore, rejecting/validating this hypothesis is the first and most important step of the procedure.

2 Data acquisition and pre-processing

In this chapter, we provide an overview of the methods followed for the acquisition of the audio data and instrumental gestures, as well as the computational tools used to perform an alignment between the score and the performance of each musician.

2.1 Recordings Procedure

Several recording sessions were carried out involving motion capture using an EMF commercial device (as detailed in [5]), piezoelectric pick-up microphones fitted on the bridge of each violin, and ambient microphones capturing the overall produced sound. For the moment, these recordings focus on the scenario of two interacting violinists in different experimental set-ups.

The first round of experiments featured two professional violinists recording excerpts from W. A. Mozart's *12 Duets* (KV 487), J. S. Bach's *Concerto for Two Violins* (BWV 1043) and L. Berio's *Duetti per due violini*. The experimental set-ups consisted of:

- a *solo* set-up, where each musician performed their part alone
- a *normal* set-up, where they performed their respective parts together as in a normal duet situation, and finally
- a *switched score* set-up where they performed together, however with the scores switched as in respect to the normal set-up.

A second round of experiments was carried out shortly afterwards involving two amateur violinists, and using the same audio and motion capturing techniques. The subjects were tasked with playing together for the first time pieces of very basic difficulty which they were not familiar with. The experimental set-ups for the second

round consisted of the same set-ups as the previous experiment; however, the pieces were (as mentioned above) simpler, including performing a popular and well-known melody (*Greensleeves*) in unison, and performing one of the duets by L. Berio at a steady, slow tempo.

2.2 Score-performance Alignment

In order to measure the deviations between the recorded performances and their respective scores, it was necessary to perform a score-performance alignment in the temporal domain.

However, aligning an audio recording to a score is, given the current state of the art in signal processing, a difficult task. This is even more the case for a continuous-excitation instrument such as the violin where the note attacks are varied and smooth; for this reason, we exploited the recorded motion capture data and its derived performance descriptors (bow position, bow transversal velocity, bow pressing force etc.) using the method described in [5]. In this method, bow direction changes as well as more subtle measurements such as an estimation of the applied bow force provide the most probable candidates for note changes, combined with information extracted from the audio (such as the fundamental frequency curve and the root mean square energy of the recorded sound). These features are given as input to an implementation of the Viterbi algorithm, which calculates the temporal alignment between the score and the recorded performance.

3 Towards an Analysis of Intonation Adjustments

In this first phase of this work, our main direction was to study the mechanism through which the violinists adjust to each other's pitch; for the experimental set-ups detailed above, to determine if violin 1 adjusts to violin 2 or the opposite, and how (timing of intonation changes, critical band of acceptable pitch difference, et cetera).

3.1 Temporal matching of different experimental set-ups

Since the recordings in the experimental set-ups were performed without a metronome, it was necessary to time-warp the performances in order to compare pitch deviations between violinists 1 and 2; in the *solo* recordings, for example, this comparison is impossible since the two recordings of violinists 1 and 2 were not temporally synchronized.

This was achieved by applying a note-by-note temporal warping algorithm based on resampling the signal between note onsets and restoring its original pitch using an implementation of the phase vocoder algorithm. Besides producing an accurate and non-destructive temporal alignment, this approach can be also very useful in performing user evaluation tests, where subjects can hear the *normal* duet recordings and the *solo* aligned-duet recordings and rate the quality of the duet's intonation; thus

investigating whether intonation adjustments alone (i.e. without temporal mismatch) can provide enough information to discriminate between solo and duet recordings.

3.2 Data post-processing and score representation

All data analysis takes part in MATLAB. Fundamental frequency (pitch) estimation is performed with the YIN [6] algorithm, while the score is imported as a MusicXML file, and converted to a time series. All time series are resampled to 1KHz, to facilitate comparisons between time series; finally, pitch estimation errors (such as octave errors) are removed with the use of pitch guides.

Figure 1 presents an excerpt of the J.S. Bach concerto recorded in the first experimental set-up.

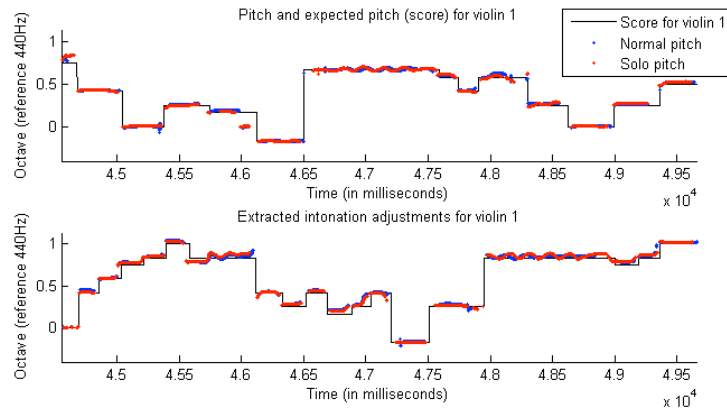


Fig. 1. Expected pitch (score) [black], recorded pitch in the *normal* set-up [blue], and recorded pitch in the *solo* set-up [red], for an excerpt of J.S. Bach’s ‘Concerto for two violins’.

4 Preliminary results & discussion

Some preliminary results and their implications are already visible from the data processing step. Namely, as it can be observed in Figure 1, the professional musicians employed for the first round of experiments were able to reproduce the same intonation (seen in fig.1 as pitch curves) with impressive accuracy; the mean difference (in pitch cents) between the *solo* and *normal* recordings, in regard to violinist 1 and 2, can be seen in table 1:

Table 1. Mean difference and standard deviation (in pitch cents) between *solo* and *normal*.

Piece	Mean difference	Standard deviation
Bach, violinist 1 (professional)	-4.976 cents	25.4 cents
Bach, violinist 2 (professional)	3.149 cents	32.7 cents

Berio n.17, violinist 1 (professional)	-8.165 cents	31.1 cents
Berio n.17, violinist 2 (professional)	-1.404 cents	13.7 cents
Berio n.24, violinist 1 (amateur)	0.610 cents	15.0 cents
Berio n.24, violinist 2 (amateur)	-2.813 cents	12.5 cents
Greensleeves, violinist 1 (amateur)	-4.659 cents	18.8 cents
Greensleeves, violinist 2 (amateur)	-1.340 cents	14.35 cents

If we consider Rossing’s [7] estimation that the JND (just-noticeable-difference) for pitch is ~ 5 cents, it is immediately evident that whichever intonation adjustments take place, they are in response to very subtle deviations from the expected pitch. Moreover, these adjustments are also partially obscured by several performance aspects, such as vibratos, glissandi between notes, and of course the shortcomings of pitch estimation algorithms.

To demonstrate this concept more clearly, Figure 2 shows a scatter plot between the pitch deviations of violinists 1 and 2, for the *normal* and *solo* set-ups.

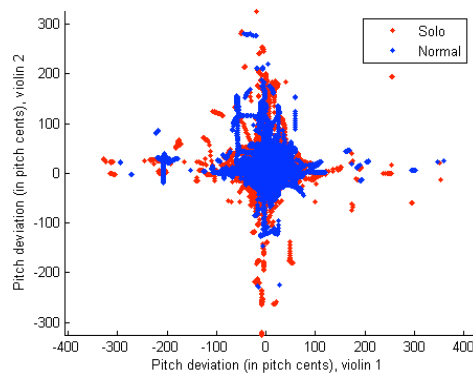


Fig. 2. Pitch deviation for violinist 1 versus pitch deviation for violinist 2, for an excerpt of J.S. Bach’s ‘Concerto for two violins’.

As it can be seen in the above figure, there is no visible correlation between the pitch deviations of the two violinists, and there is little (if any) visible difference between the *solo* and *normal* set-ups. The implications of such an observation lead us to the conclusion that, whichever synchronization phenomenon occurs between the intonation adjustments in a duet, it is not consistent throughout the piece, and cannot be viewed as a process which is either invariant to time or cyclostationary.

To this end, our next step has been to employ several dependence measures between the two time series, such as Mutual Information, Granger Causality and non-linear directional coupling detection algorithms derived from computational neuroscience [8].

The latter coupling measure used (referred to as the *L*-measure) is capable of providing also the *directionality* of the interdependence, in our case the strength with which violinist 1 influences violinist 2 and vice versa; figure 3 displays some preliminary results based on this measure, where we were able to observe an increase of average coupling strength in the *normal* set-up compared to the *solo* set-up. The

most evident separation between the two set-ups can be found in the amateur musicians' recordings - since the second violinist was less adept in *prima vista*, the first violinist had to adapt quite a lot in order to balance the performance.

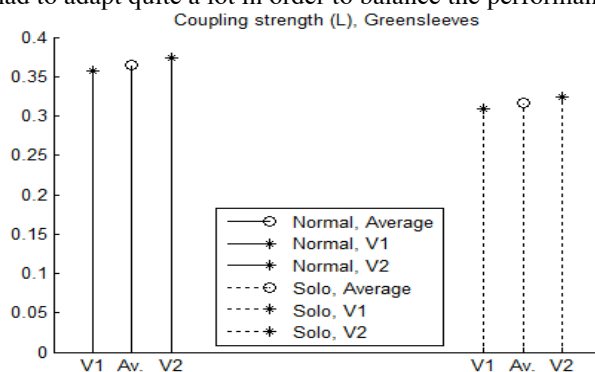


Fig. 3. Coupling strength between the pitch deviations of violinists 1 and 2, for the *normal* and *solo* recordings of one of Greensleeves (amateur musicians). The x-axis shows three different coupling strengths for every recording: V1 \Rightarrow V2, V2 \Rightarrow V1 and average coupling strength.

However, before we can extract clear conclusions from the use of these measures, a refinement is required in the initial features used for this approach; such an effort is currently underway.

Acknowledgements

The work presented on this document has been partially supported by the EU-FP7 ICT SIEMPRE project.

References

1. Widmer, G., & Goebel, W. : Computational Models of Expressive Music Performance: The State of the Art. *Journal of New Music Research*, 33(3), 203-216.
2. Keller, P. E. : Joint action in music performance. *Emerging Communication*, 10, 205. IOS press.
3. Moore, G. P., & Chen, J. : Timings and interactions of skilled musicians. *Biological cybernetics*, 103(5), 401-14. (2007)
4. Kalin, G. : Formant Frequency Adjustment In Barbershop Quartet Singing. Doctoral dissertation, KTH, Department of Speech, Music and Hearing. (2005)
5. Maestre, E. : Modeling instrumental gestures: an analysis/synthesis framework for violin bowing. Doctoral dissertation, Univ. Pompeu Fabra, Barcelona, Spain. (2009)
6. De Cheveigné, A., & Kawahara, H. : YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111(4), 1917-1930. (2002)
7. Rossing, T.D. : *The Science of Sound* (2nd Edition), Addison-Wesley (1990)
8. Chicharro D., Andrzejak R.G. : Reliable detection of directional couplings using rank statistics. *Physical Review E*, 80, 026217 (2009)