

Automatic characterization of ornamentation from bassoon recordings for expressive synthesis

Montserrat Puiggròs, Emilia Gómez, Rafael Ramírez, Xavier Serra

Music technology Group
Universitat Pompeu Fabra
Barcelona, Spain

{mpuigros, egomez, rramirez, xserra}@iua.upf.edu

Roberto Bresin

Speech, music and hearing
Royal Institute of Technology
Stockholm, Sweden
roberto@kth.se

ABSTRACT

Expressive performance characterization is traditionally based on the analysis of the main differences between performances, players, playing styles and emotional intentions. This work addresses the characterization of expressive bassoon ornaments by analyzing audio recordings played by a professional bassoonist. This characterization is then used to generate expressive ornaments from symbolic representations by means of Machine Learning.

INTRODUCTION

Expressive performance characterization analyzes differences in performances, performers, playing styles and emotional intentions (Juslin and Sloboda 2002). Most research focus on studying timing deviations, dynamics and vibrato (see for instance (Sudberg et al. 2003) and (Bressin and Friberg 2000)). Nevertheless, there is less research devoted to ornamentation. Ornaments are indicated in the score, without any explicit information about timing and dynamics. Some works have already studied the behaviour of ornaments from piano performances (Moore 1992). We study here how this study for the piano can be extended to other instruments, as the bassoon, a woodwind instrument. Due to the unavailability of expressive MIDI extracted from bassoon performances, we analyze directly expressive audio recordings played by a professional musician.

METHOD

The block diagram of the system is presented in Figure 1. We divide this study in two main stages, analysis and synthesis, which correspond with the main goals of this work. First, to study the behaviour of ornamentation by analyzing timing and dynamics from bassoon recordings. Then, the acquired knowledge is used for the generation of expressive trills in symbolic notation using some machine learning tools.

In the analysis stage, we describe the process to describe ornament's behaviour, more precisely trills and appoggiaturas, by means of automatically extracting timing and dynamics information from bassoon recordings. The re-

cordings used in this study belong to a Sonata of Michel Corrette (composer of XVIII century). Each movement is played in three different tempi, obtaining a total of 96 ornaments including trills and appoggiaturas. The result of this analysis is a melodic description for each ornament.

In the synthesis stage, we first study the ornament's behaviour using different machine learning methods from the information obtained in the analysis.

Finally, and also by another machine learning method, we generate expressive ornaments in symbolic notation, introducing them as notes in the input melody.

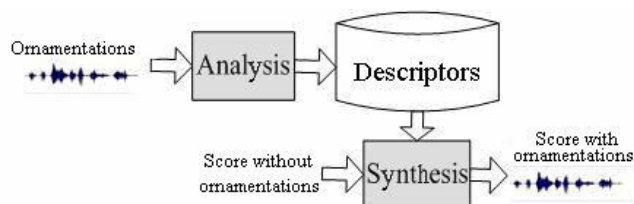


Figure 1: Block diagram of the system

Analysis

The analysis stage consists on the melodic description of sound material. As mentioned above, we characterize a set of expressive recordings of a Sonata by Michel Corrette (a baroque epoch's sonata) played by a professional bassoon performer. There are three movements: "Adagio", "Allegro moderato" and "Affettuoso". Each movement is played in three different tempi. "Adagio" is played at 50, 68, 100 bpm and "Allegro moderato" and "Affettuoso" is played at 60, 92, 120 bpm.

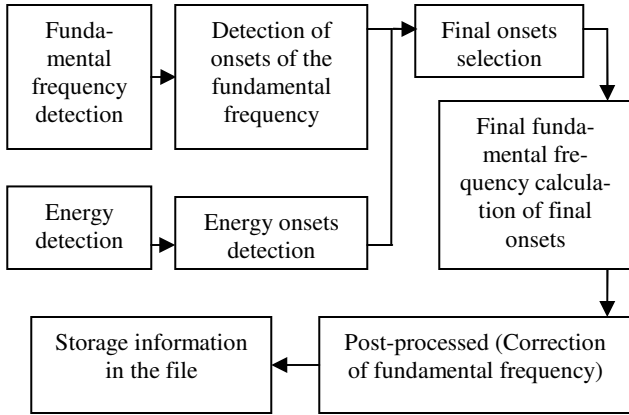


Figure 2: Description of the steps of analysis

Thus we have obtained a total of 96 ornaments (trills and appoggiaturas), as a collection to study the different expressive variations from the same ornaments.

The analysis is carried out by the algorithm shown in Figure 2. Some of the steps have already been presented in (Gomez 2002, Gomez et al. 2003). We have adapted the algorithm parameters to the specific characteristics of the bassoon in order to consider pitch range, note duration (between 0.05 and 0.04 seconds as trill’s execution is very quickly) and short intervals between notes, 1 or 2 semitones. We first estimate the instantaneous (on a frame basis) fundamental frequency and energy from the audio recordings, only analysing the ornaments obtained from these interpretations. After this computation we compute a perform a segmentation in order to obtain onset, offset and fundamental frequency information for each ornamental note. The onset detection algorithm is based on (Klapuri 1999). We can see an example in Figure 3.

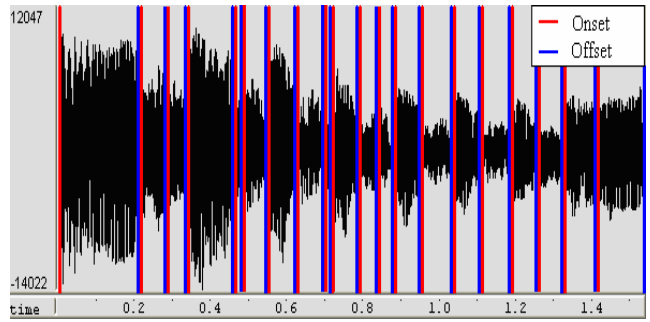


Figure 3: Onsets and offsets detected from instantaneous energy and fundamental frequency. The red lines indicate the onsets and the blue lines the offsets.

After detecting all possible onsets, we make a selection of onsets choosing the most suitable ones through a set of rules. First we verify that notes are consecutive, i.e. there is no overlap between them. When there is an overlap, we have to move the offset in order to make it equal to the next onset, as in Figure 4 and.

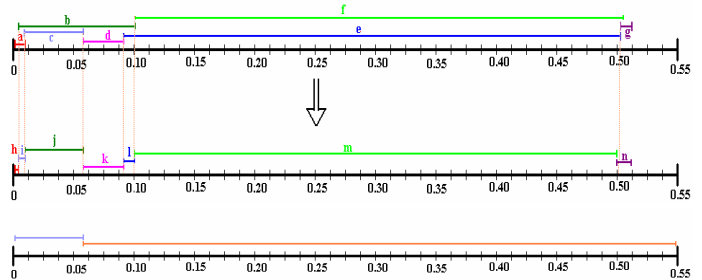


Figure 4: Correction of detected onsets. The top panel shows the estimated onsets. In the middle panel, overlapped notes have been merged, and in the bottom panel, too-shorts are also deleted.

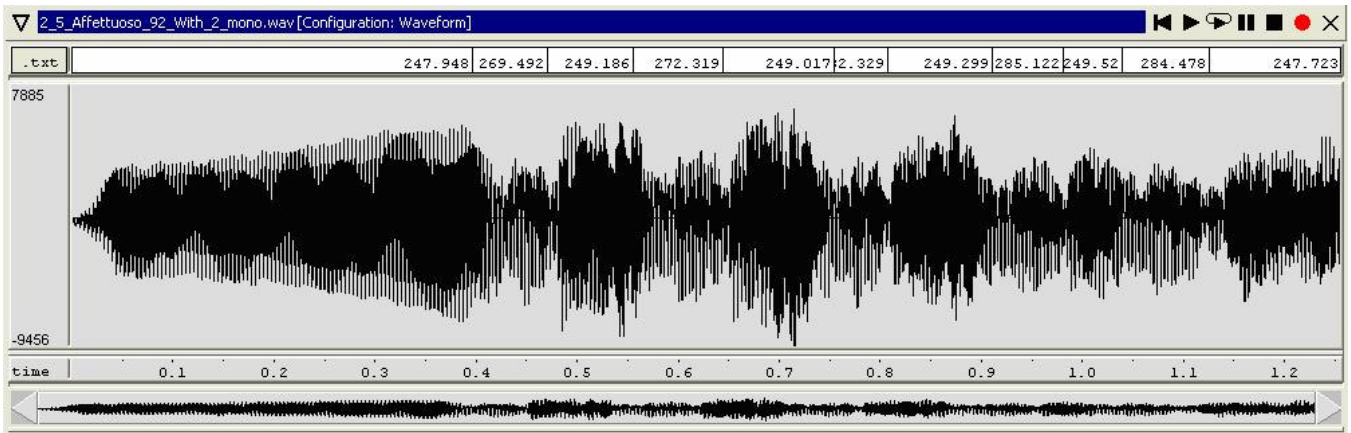


Figure 5: Example of the analysis results.

Having the final onset values, we compute again the fundamental frequency for each of the ornamental notes. Then, we correct fundamental frequency values in order to check the alternation of the notes of the trill and so that the distance between two notes only can be 1 or 2 semi-tones.

Hence, we have obtained the final note’s descriptors: onsets, offsets and fundamental frequency. In Figure 5 it is possible to see an example of the final result with all descriptors. We store the descriptors in a text file, as shown in Figure 6.

Although these descriptors we also save the context of each ornament: the note anterior and posterior with their respective durations, the beat, tempo and movement.

| Onset | Offset | Frequency |
|-----------|-----------|-----------|
| 0.0000000 | 0.2449990 | 392.00 |
| 0.2449990 | 0.2958233 | 419.784 |
| 0.2958233 | 0.3773240 | 392.841 |
| 0.3773240 | 0.5108390 | 419.784 |
| 0.5108390 | 0.7183670 | 398.419 |
| 0.7183670 | 0.9171880 | 354.985 |
| 0.9171880 | 1.1494100 | 392.135 |

Figure 6: Example of melodic descriptors. The onset and offset are coded in seconds and fundamental frequency in Hz. These descriptors will be used in the synthesis part.

Synthesis

The synthesis block deals with the generation of expressive ornamentations by using the results of the analysis part.

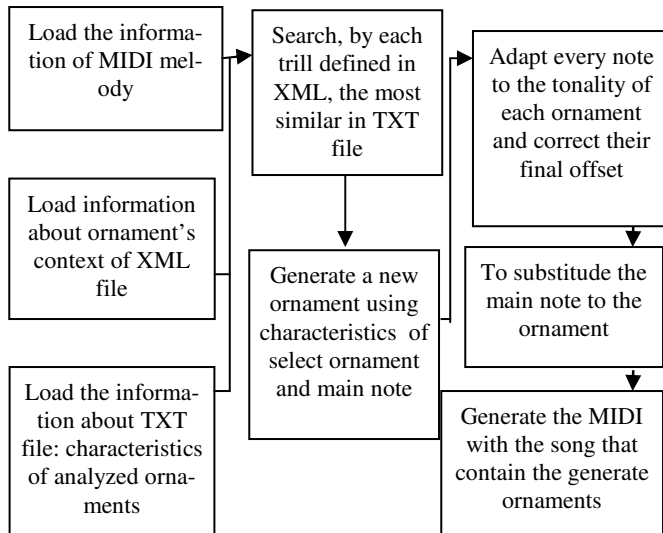


Figure 7: Description of the steps of analysis

Given a score of a melody with indicated ornaments, we define the context of each note that contains an appoggiatura or a trill, using a XML format. Information about the current note includes the note's duration, pitch and metrical position, while information about its context includes the duration of previous and following notes, extension and direction of the intervals between the note and both the previous and the subsequent note and tempo of the performance.

Once we define the context, we apply a nearest neighbour algorithm for generating the expressive ornament. The algorithm selects the most similar trill (in terms of musical context) in the training examples and adapts it to the new musical context (e.g. the key of the piece).

After finding the ornament with higher similarity, their descriptors are adapted to the characteristics of the input note, pitch and duration, and the new ornamented note is generated.

Once we have the descriptors of corresponding ornament we consider the main note’s descriptors (beginning and end time and fundamental frequency). Bearing these parameters in mind, we adapt them to the behaviour of the once already analyzed. We consider if it is an ascending or descending ornament, the duration of each note of the trill and the duration of the main note. We scale duration and fundamental frequency information, taking in a count the tonality of the new melody, and transform it into a MIDI representation. Finally, when we have the new ornament, we insert it into the symbolic representation of the new melody.

RESULTS

1.1 Statistical analysis

The melody estimation has been successfully adapted to the particular analysis of bassoon ornaments. The statistical analysis of the duration of the ornamental notes reveals a similar behaviour to previous studies on piano (Brown, Judith. 2003). The speed of execution is around 8 notes per second for most of the trills. Figure 8 shows the distribution of the notes classified in the three movements of the analyze piece: Allegro, Affettuoso and Adagio. We can observe that majority group is of 8 notes, as mentioned below.

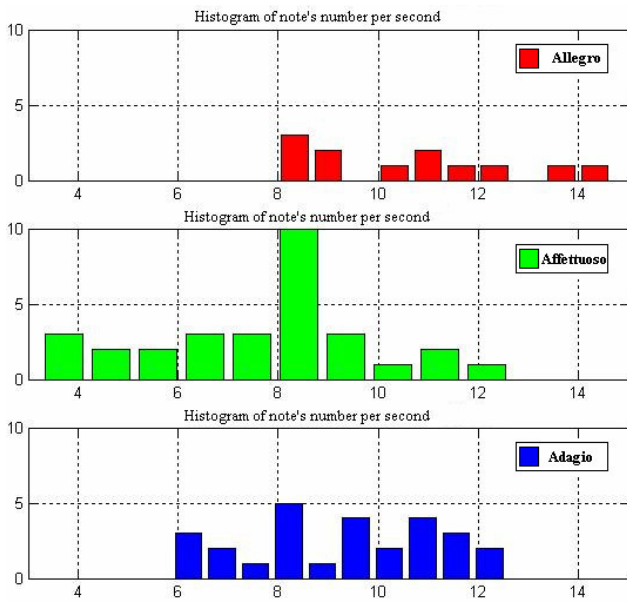


Figure 8: Distribution of number of notes per second for the three movements: Allegro, Affettuoso and Adagio.

Another interesting result is that we can clearly distinguish two groups of trills. In the first group, that of the slow tempi (notes with long duration), there is a difference among both extreme notes (the initial and the final note) and middle notes. The first and the last note are usually longer than the central ones, as shown in Figure 9. In the second group, for fast tempi (short notes), trills are usually converted into appoggiaturas, as shown in Figure 10.

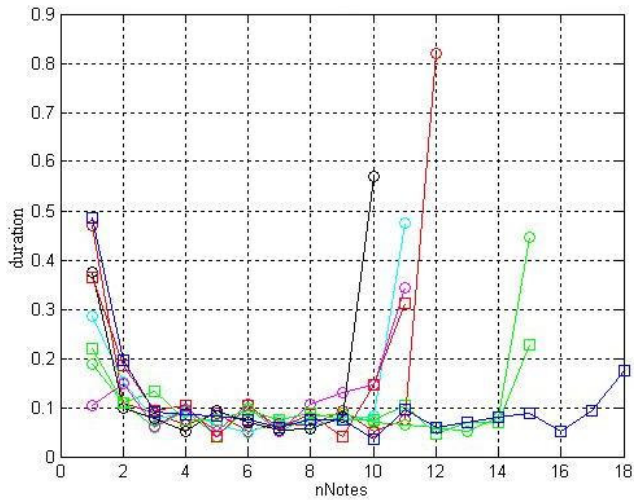


Figure 9: Duration of ornamental notes for the ten longest trills. We observe that the first and last note have a longer duration than the rest.

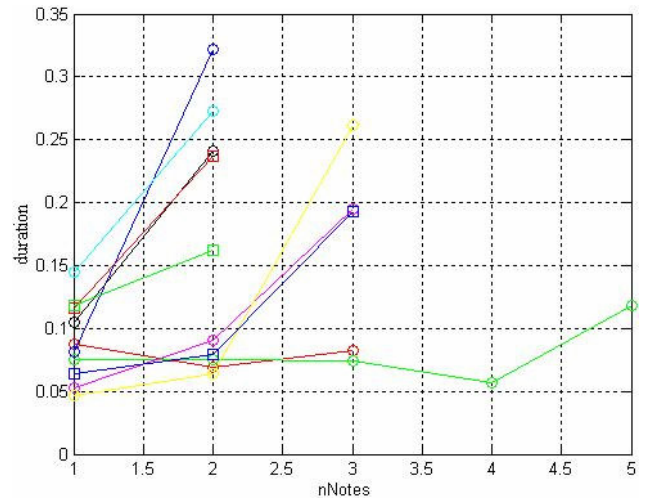


Figure 10: Duration of ornamental notes for the ten longest trills. We observe that the behaviour is the same that for an appoggiatura. The first note is shorter than the second, which acts as the main note.

Finally we can sometimes identify some regularity in the execution of central notes duration. In this situation, we can speak about “controlled” trills, as opposite to “non-controlled” trills. In Figure 11 we show an example of a controlled trill.

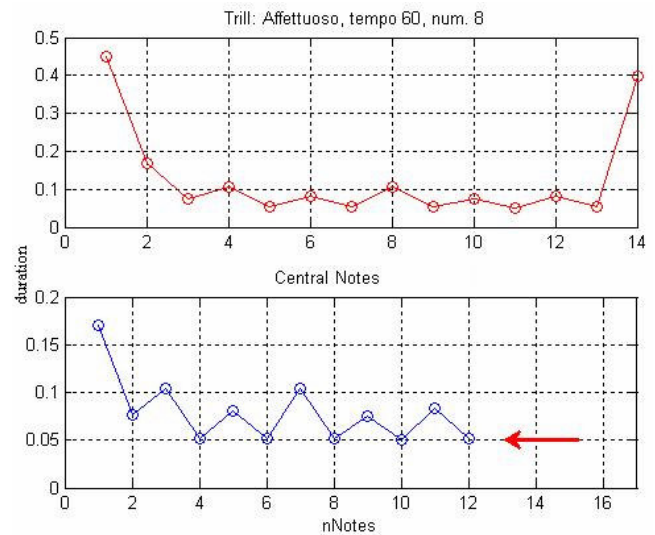


Figure 11: Evolution of note duration for a “controlled” trill. The central notes are played with regularity.

Generation of ornaments

Figure 13 and Figure 14 show an example of ornaments generated with this method.

Les Delices de la Solitude

Michel Corrette
(1709-1795)

Affettuoso $\text{♩} = 92$

Bassoon

Figure 13: Original score without trills. This is a fragment of a bassoon melody of Affettuoso movement and tempo 92 pulsations per second.

Les Delices de la Solitude

Michel Corrette
(1709-1795)

Affettuoso $\text{♩} = 92$

Figure 14: Final score with generated ornaments, They are indicated with a red line.

CONCLUSIONS

This study presents an approach for the automatic analysis and generation of expressive ornaments of bassoon using automatic melodic description and machine learning techniques.

There seems to be regularities on the trills if we distinguish two groups for long and short trills. Our results agree with previous studies for piano, although it seems to be easier to perform trills in bassoon, because it is softly to play than piano. Ultimately we can reproduce the behaviour in a MIDI synthesizer.

Further work is centred in increasing the analyzed collection in order to obtain a robust model and to extent it to other musical instruments.

REFERENCES

- Bresin, R. & Friberg, A. (2000) Emotional Coloring of Computer-Controlled Music Performances. *Computer Music Journal*, 24(4), pp. 44-63
- Brown, Judith C. (2003). Independent component analysis for automatic note extraction from musical trills, *Journal of the Acoustic Society of America*, 115, pp. 2295-2306.

Gómez E. (2002). *Melodic description of audio signals for music content processing*, PhD predoctoral Thesis, Universitat Pompeu Fabra.

Gómez, E. Grachten, M. Amatriain, X. Arcos, J. (2003). *Melodic characterization of monophonic recordings for expressive tempo transformations*, Proceedings of Stockholm Music Acoustics Conference 2003; Stockholm, Sweden

Juslin, Patrik N., Sloboda, John A. (2002). *Music and emotion*, Oxford University Press.

Klapuri, A. (1999). *Sound Onset Detection by Applying Psychoacoustic Knowledge*. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP.

Moore, G. P. (1992). *Piano trills*, *Music Perception* 9(3), pp. 351-359.

Sundberg J, Friberg A, and Bresin R (2003) *Attempts to reproduce a pianist's expressive timing with Director Musices performance rules*, *Journal of New Music Research*, 32(3), pp. 317-326.