CONTENT ANALYSIS OF NOTE TRANSITIONS IN MUSIC PERFORMANCE

Loureiro, Mauricio
Escola de Música
UFMG
mauriciloureiro@ufmg.br

Yehia, Hani
Escola de Engenharia
UFMG
hani@ufmg.br

de Paula, Hugo
Dep. de Informática
PUC - Minas
hugodepaula@gmx.net

Campolina, Thiago
Escola de Engenharia
UFMG
thicampolina@yahoo.com.br

Mota, Davi
Escola de Música
UFMG
davialvesmota@yahoo.com.br

ABSTRACT

Different aspects of music performance have been quantified by a set of descriptor parameters, which try to incorporate the resources used by the performer to communicate his/her intention of expressiveness and intelligibility. The quality of note transitions are quite important in the construction of an interpretation. They are manipulated by the performer by controlling note durations and the quality of attacks and note groupings. These characteristics can be modeled by parameters that may describe what happens between the notes of a musical sentence, which attempt to represent how we perceive note articulations and groupings of legato or detached notes. On the other hand, the quality of transitions between legato notes may be related to the musician's abilities, to the reverberation characteristics of the performance room and to the acoustic characteristics of the instrument. This text illustrates methods of extraction and definition of descriptor parameters related to the quality of transitions between notes, which are capable to reveal relevant aspects about the accomplishment of these transitions. The procedures here described integrate a model of analysis of the expressiveness of performance in monophonic musical instruments. The samples used consist of recordings of interpretations of excerpts of the classic repertoire for solo clarinet.

1. INTRODUCTION

The comprehension of the processes involved in the production and perception of an expressive performance has been approached by models that quantify the player's expressive intentions upon acoustic information. Studies on musical expressiveness have demonstrated that musicians use small variations of duration, articulation, intensity, pitch and timbre to communicate to the listener aspects of the music that they interpret [2, 3]. By comparing performances of different musicians as well as different interpretations by the same musician, these deviations can be perceived with a surprising clarity, even by non specialized listeners. The quantification of the interpreter's expressive intentions upon such deviations involves the identification and measurement of a set of descriptor parameters defined and calculated from information extracted from the audio signal of the recorded musical performance. In order to incorporate different aspects of the resources used by the performer to communicate his/her intention of expressiveness and intelligibility, these parameters may be established for different segmentation levels, such as musical notes, groups of notes, specific areas in the extent of a same note or in the transition between consecutive notes.

SMC 2009, July 23-25, Porto, Portugal

1.1. Transitions between Consecutive Notes

Characteristics related to the quality of the transitions between consecutive notes are decisive in the construction of an interpretation. They are manipulated by the performer by controlling note durations and the quality of attacks and note groupings. These characteristics can be modeled by parameters that try to describe the ephemeral sonorities that happen between the notes of a musical sentence, as an attempt to represent how we perceive note articulations and groupings of legato or detached notes. The quality of transitions between notes may be related not only to the performer's intention, but also to his/her musical skills, to reverberation characteristics of the performance room and to acoustic characteristics of the instrument.

During the short period of time of a transition between two consecutive notes of a musical phrase, inharmonic sounds or strange frequencies to both of the involved notes can occur. On the clarinet, occurrences of lower frequencies in low levels of intensity are very common in passages of fast and long range leaps or between different registers. This kind of sonority was explored by Carl Maria von Weber in the Minuet of his *Grand Quintetto* in B flat Major op. 34 for clarinet and string quartet. The clarinet plays a three-note arpeggio, in which a leap from B 4[1] (the instrument's longest vibration mode), to A 5 (vibration mode confined in a small portion of the tube)

[1] As written for B flat clarinet.

produces an F 4, that seems to be a sub-harmonic of A 5. While this is not well perceived as a pitch, this three-note arpeggio enhances the color of a dominant seventh chord (G-B-D-F), in which the A 5 on the clarinet works as a passing tone. Weber repeats the same arpeggio transposed up a minor third, a passage acoustically similar and with the same characteristic sonority. Did Weber hear this "sub-tone" as a the seventh of this dominant chord (Figure 1)?

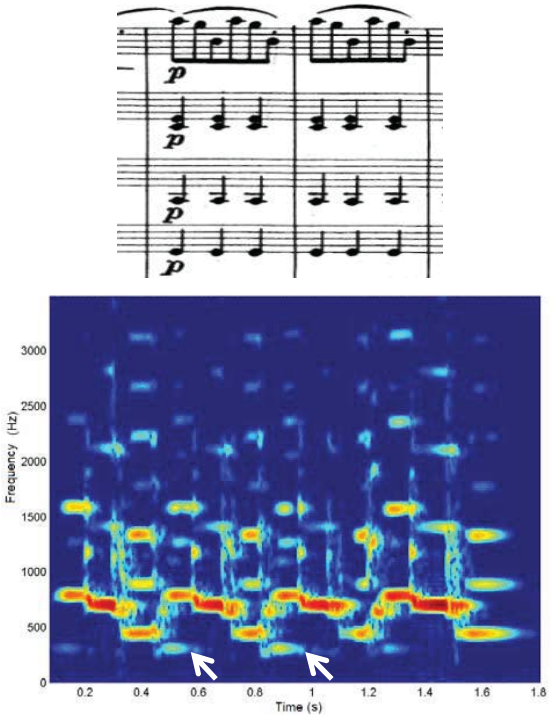

Figure 1. Sonogram of three-note arpeggio of the Minuet of the *Grand Quintetto* in B flat Major op. 34 for clarinet and string quartet by Carl Maria von Weber, showing a sub-harmonic pitched F during the leap from B 4[1] to A 5 (indicated by arrows).

Such sonorities can be found quite often in the clarinet repertoire. Bohuslav Martinu in his *Sonatina for Clarinet and Piano* and Alban Berg in the last of the *Four Pieces for Clarinet and Piano* Op. 5, explored the sound of fast passages between the lower and mid registers of the instrument. In *Clarinet Threads* for clarinet and tape, Denis Smalley explores the spectral morphologies of these sonorities by asking the player to reduce air and lips pressure. These sub-tones, which constitute the thematic material of the work, can only be produced in very low intensity levels. The balance of dynamic level is guaranteed by electronic amplification of these sounds.

[1] As written for B flat clarinet.

This text illustrates methods of extraction and definition of descriptor parameters related to the quality of transitions between consecutive notes, that may describe not only the degree of the musician's intentional manipulation of note articulations, but also aspects related to the musician's ability in accomplishing legato transitions and to acoustic characteristics of the instrument. The procedures here described integrate a model of analysis of the expressiveness of performance in monophonic musical instruments. The used samples consist of recordings of interpretations of excerpts of two major pieces from the traditional classic repertoire for solo clarinet.

2. METHODS

2.1. Materials

The compositions used in this study were the *Quintet for Clarinet and String Quartet* in A Major Kv 588 and the *Concert for Clarinet and Orchestra* in A Major Kv 622, both by W. A. Mozart.

Different solo performances of excerpts from these pieces were recorded without the accompaniment, in the same day, under the same conditions of acoustic environment and equipment. Four different professional clarinetists, 2 of them members of the Philharmonic Orchestra of …(my state), performed each excerpt 3 times, in different ways, according to the instructions:

- Performance **P1** – an expressive performance, as in a concert situation, in which the piece would be played entirely.
- Performance **P2** - without any expression intention, but trying to adopt the same tempo of **P1**.
- Performance **P3** - an expressive performance, as in the first execution, however with different expressive intentions.

2.2. Segmentation

The first step for an appropriate estimate of descriptor parameters is to segment the signal into events to be analyzed, such as musical notes, note groups, or specific regions within a single note. Segmentation is not a trivial problem, even on monophonic musical signals, especially if we consider the subjectivity in the discrimination of these events. Note onsets and offsets were detected on the RMS envelope averaged for 23 ms using an adaptive threshold, as suggested by De Poli [1], calculated as the average energy in a certain neighborhood (1 s for a step of 6 ms) of each point of the RMS. Onset and offsets are detected by searching the minimum RMS between two consecutive values crossed by this dynamic threshold. The

estimate of the fundamental frequency changes, with a pitch threshold below a half tone, helped the segmentation in cases where the detection of onsets and offsets was not possible by means of energy level only, such as legato notes.

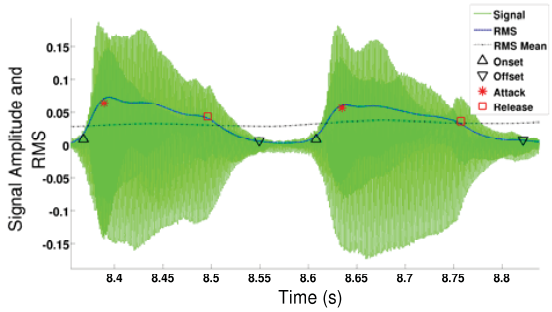

Figure 2. Detection of note onsets and offsets on the RMS envelope using a dynamic threshold calculated as the average energy surrounding each point of the RMS (1 s for a step of 6 ms). Onset and offsets are detected by searching the minimum RMS between two consecutive values crossed by this threshold.

The end of attack was defined as the first amplitude maximum after the note onset, and the beginning of release, as the first amplitude maximum before the note offset. These points were detected by searching for maximum variations of the first derivative of the RMS signal. It doesn't exist in the literature a measurement method that can describe the attack unequivocally [4]. This definition of attack is adequate to describe the attack in most situations, but further consideration was necessary in cases where maximum amplitude was reached much later in the sustained segment of the note.

The presence of transients during note transitions makes possible the use of spectral flux to detect the end of attack, since this point can be related to the reestablishment of harmonics amplitude correlation after note transitions:

$$SF = \frac{1}{T}\sum_{p=1}^{T}\left|r_{p,p-1}\right| \qquad (1)$$

where $r_{p,p-1}$ is the correlation coefficient among spectral amplitudes calculated at instants t_p and t_{p-1} of the signal of duration T.

The spectral flux confirmed most of the end of attack estimated by energy local maxima and was also able to detect these points where the energy method was not.

2.3. Descriptors for Note Transitions

In order to analyze the quality of the transition between consecutive notes two descriptors were defined: *articulation index* and *legato index*. Some aspects of the quality of note transitions are related to the performer's intentional manipulations of note articulations, which are well controlled by the musician by manipulating note durations and attack quality. Hence, articulation is considered to be closely related to the performer's intentions of expressiveness and intelligibility.

The *articulation index*, defined as the ratio between note duration DR (time interval between note onset and offset) and the time interval between both note onsets (known as intra-onset-interval - IOI), was used to describe the intentional manipulation of note duration, by the performer:

$$AI(i) = 1 - \frac{DR(i)}{IOI(i)} \qquad (2)$$

This index is appropriate to describe transitions between detached notes, usually produced in the clarinet with abrupt interruptions of the air flow by slightly beating the tongue on the reed. With this action, the player controls the quality of the attacks as well as the duration of each note. It should be pointed that the performer's ability to control note duration depends closely on the ambience reverberation conditions.

On the other hand, this index is not adequate for describing the legato transitions, since it assumes values close to 0 for most legato notes.

2.4. Legato Index

Legato notes on the clarinet are produced by means of a single blow without interrupting the air flow during the transition. To investigate the quality of transitions between notes played with the intention to be legato, we used a descriptor suggested by Maestre [4], defined as a comparison to an ideal legato between 2 notes without any decrease of energy, represented by the straight line traced from the beginning of release of a note to the end of attack of the subsequent note. The index is calculated by means of the area A_1 between this line and the energy curve and the total area $A_1 + A_2$ below this straight line:

$$LI = 1 - \left[\frac{A_1}{A_1 + A_2}\right] = 1 - \left[\frac{\sum_{t=release(i)}^{attack(i+1)}(r(t) - RMS(t))}{\sum_{t=release(i)}^{attack(i+1)} r(t)}\right] \qquad (3)$$

r being the line traced from the beginning of release of a note to the end of attack of the subsequent note.

This index appears to be related to the musician's abilities, to the ambience reverberation conditions and to the acoustic characteristics of the instrument.

3. DISCUSSION

3.1. Articulation Index

The capability of these two descriptors to represent the note transitions of a musical phrase played in a monophonic instrument is illustrated below with the first sentence of the main theme of the first movement of the Quintet for Clarinet and String Quartet in A Major Kv 588 by W. A. Mozart, bars 118 through 124 (Figure 3).

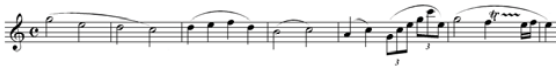

Figure 3. First sentence of the main theme of the first movement of the Quintet for Clarinet and String Quartet in A Major Kv 588 by W. A. Mozart (bars 118 through 124).

Figure 4 shows the evolution of the articulation index of the note transitions along performances **P1** and **P2** of the first five bars of the phrase of Figure 3, by one of the clarinetists. It can be observed that all transitions up to 13 (between the first 2 eighth triplets G-C of the 5th bar) present index values below 10 % for both performances. This is because all these notes were played legato, with the exception of some soft articulations. Transitions 14 through 18 of performance **P1** presented index values varying from 15 % to 26 %. In this performance, all these notes were articulated**,** while in performance **P2** the triplets arpeggio was played legato up to the high C (adopted articulation is shown in the upper part of Figure 4).

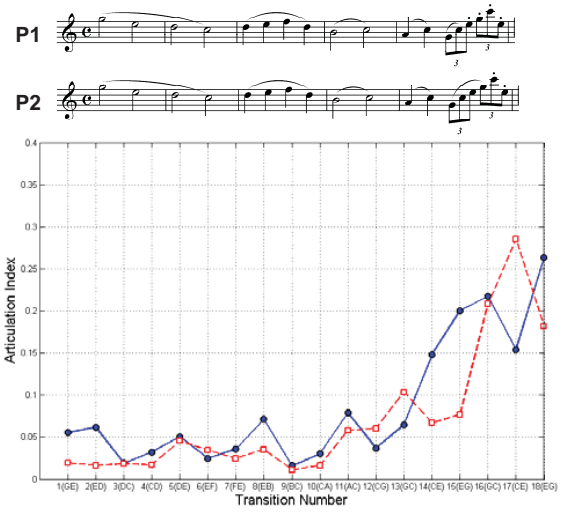

Figure 4. Articulation index of performances **P1** (continuous line) and **P2** (traced line) of the first 5 bars of the main theme of the first movement of Mozart Clarinet Quintet Kv 588, by the same musician (adopted articulation shown on the top).

As expected, this index was not adequate for describing transitions between legato notes, as the smooth articulations performers use to define phrase boundaries, such as transition 8 (D of the end of third bar to B of the beginning of the fourth bar). Despite the perceptual evidence of the articulation produced in performance **P1** in this transition, the index scored only 7 %, which appears to show that this index is not adequate for describing quasi legato articulations.

3.2. Legato Index

Figure 5 shows the legato index of performances **P2** (musician instructed to play with no expression) and **P3** of the first 4 bars of the same excerpt and by the same clarinetist of Figure 4. Lower values of this index confirm the phrase articulations at transitions 4 and 8, evident in both performances. Performance **P3** presented over all lower values than performance **P2.** This might be explained by consistent energy variation along the whole sequence of legato notes in performance **P3** (played with expression), which was not only clearly audible, but also confirmed by the RMS curve. Although it is well known that a good legato on the clarinet is achieved with a very uniform blow with minimal fluctuation of the air pressure, consistent intensity variation, as heard in performance **P3**, is not rare in "molto expressive" performances, in which the varying air pressure compromises the legato quality.

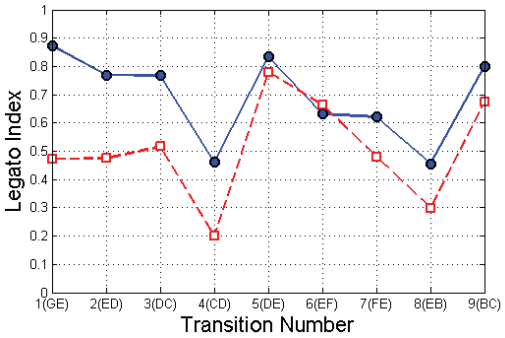

Figure 5. Legato index of performances **P2** (continuous line) and **P3** (traced line) of the first 4 bars of the main theme of the first movement of Mozart Clarinet Quintet Kv 588, by the same musician.

A sudden increase of the legato quality could be observed at transitions 5 and 9, which come right after both articulations phrase at transitions 4 and 8. Such increase, observed in several alike passages, might be explained by the fact that right after an interruption of the air flow to produce the articulation, it seems to be easier to build a more uniform air pressure for the subsequent note transition. Figure 6 shows another example of this at

transition 5 of both performances **P2** and **P3** of the same excerpt by another clarinetist. Note that this player decided not to articulate transition 8, despite the natural articulation of this phrase, marked by harmony change to the dominant.

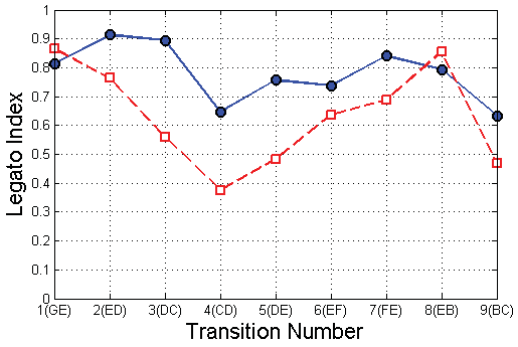

Figure 6. Legato index of performances **P2** (continuous line) and **P3** (traced line) of the first 4 bars of the main theme of the first movement of Mozart Clarinet Quintet Kv 588, by a second clarinetist.

Figure 7 illustrates another example of such behaviors of this index in fast note transitions (6/8th measure at 96 BPM): performances **P1** and **P2** of the first sentence (6th to 13th notes) of the third movement of Mozart Clarinet Concerto Kv 622, played by the second clarinetist. Better legato quality could be also observed in the non expressive performance, as well as at transition 7, which comes right after the phrase articulation of transition 6.

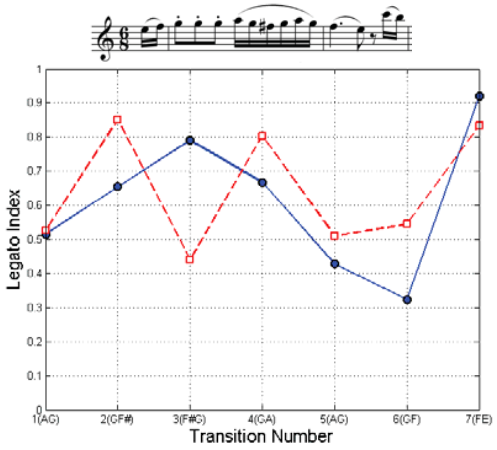

Figure 7. Legato index of fast note transitions: Performances **P1** (continuous line) and **P2** (traced line) of notes 6th through 13th of the third movement of Mozart clarinet Concerto, played by the second clarinetist (6 / 8th measure at 96 BPM).

4. CONCLUSION

This paper presented quantitative ways (namely, articulation index and legato index) to measure and compare transitions between notes, as performed by four clarinetists in three different ways. The articulation index allowed a description of the degree of the musician's intentional manipulation in the accomplishment of articulation between notes.

Measures of the legato index made possible to infer about the quality of transitions between notes executed with the intention to be legato. These results suggest a dependence of the value of this index to aspects related to the musician's ability to play legato notes as well as to acoustic characteristics of the instrument, as it corroborates many aspects of practical experiences in playing and perceiving legato notes.

Higher variance of the legato index was evident for all expressive performance, when compared with non expressive performances. However no single pattern could be derived from these variations. Possible correlation between this index with the degree of some aspect of musical expressiveness, may suggest the adequacy of this descriptor for music expression investigations.

5. ACKNOWLEDGEMENTS

The accomplishment of this work was possible thanks to the participant clarinetists. This work was supported in part by FAPEMIG (Research Funding Agency of Minas Gerais State, Brazil) and CNPq (Brazilian National Council for Scientific and Technological Development).

6. REFERENCES

[1] De Poli, G. and L. Mion, "From audio to content," In *Umpublished book*, G. De Poli, Ed. Padova: Dipartimento di Ingegneria Dell'Informazione - Università degli Studi di Padova, 2006.

[2] Gabrielsson, A., "Music Performance Research at the Millenium," *Psychology of Music*, vol. 31, pp. 221-272, 2003.

[3] Juslin, P. N., "Cue utilization in communication of emotion in music performance: relating performance to perception," *Journal of Experimental Psychology: Human perception and performance*, vol. 26, pp. 1797-1813, 2000.

[4] Maestre, E. and E. Gómez, "Automatic Characterization of Dynamics and Articulation of Expressive Monophonic Recordings," in *Audio Engineering Society 118th Convention Paper*, 2005.