Multi-modal Acquisition of Performance Parameters for Analysis of Chopin's B Flat Minor Piano Sonata Finale Op.35

Jennifer MacRitchie, Nicholas J. Bailey and Graham Hair

2006-12-20

Abstract

The finale of Chopin's B flat minor piano sonata Op.35 has confounded traditional musicological approaches and there exists little serious analysis of the piece. However, there are many recorded and live performances of the work, so a fruitful approach to compensating for the lacunae is to use engineering techniques to proceed directly to the sound of the piece.

Previous approaches to performance analysis measure variables such as tempo and dynamics, and this is usually done without much consideration of the musical structure. This paper will discuss an approach which involves the collection of audio, video and MIDI data over several performances with different interpretations in performance and how they illuminate aspects of the musical structure. Collected multi-variate performance data is correlated, with care being taken that the analysis exposes the musical semantics in a quantifiable manner. The results of this are then to be compared with traditional score-based music analysis to establish a better understanding of the piece, as well as developing and innovating technology for the better representation of music by computational means.

1 Introduction - Chopin's Monophonic Finale

The finale of Chopin's B flat minor piano sonata Op.35 has been referred to as "a wild child, unique and well-nigh indescribable"[Tan05]. A short and rhythmically unrelenting finale written mostly in octaves, this binary sonata form composition has confounded traditional approaches to its analysis.

The written literature on this particular piece is very sparse and both anecdotal and impressionistic, probably because of the problematic nature of the composition. Only Charles Rosen [Ros99] has written an extensive essay and most of his statements are of the ex-cathedra type, even though his authority as a pianist prompts us to take them seriously. His comments have been taken as a first piece of secondary literature from which the present investigation is proceeding.For our purposes, this problematic nature of the work makes the data more suitable for objective, quantitative methods.

Rosen's analysis of the piece sets the first four bars as the introduction in the dominant key of B flat minor, with the chromatic main theme entering in bar 5. After bar 8, there is a transition section where the harmony of the chromaticism gradually defines the dominant of the relative major key. A new theme set in D flat major enters at bar 23 and is repeated an octave higher at bar 27. The recapitulation begins at bar 39 by literally repeating the first eight bars of the composition and then expanding the recapitulation of the following bars with parts of the transition and the second theme, moving towards a cadence.

Another similar viewpoint on the segmentation of this piece comes from Michael Talbot [Tal01].His "ground plan" of the finale is as follows:

Bars	*Key*	*Description*	*Comment*
1-8	b flat	first theme	establishing tonic
9-22	modulating	transition	chromatically unstable
23-30	d flat	second theme	diatonic
31-38	modulating	retransition	sequential progressions
39-46	b flat	first theme	reprise of bars 1-8
47-56	modulating	transition/second theme	based on bars 9-30
57-75	b flat	coda	largely diatonic

One of the first tasks of a musical analysis is to segment the data in partitions on the basis of the structural features of the work and the different characters and functions of the different sections. Crucial aspects to notice and describe empirically are points of change in various parameters such as arrival points, climax points, initiations of new musical processes and so on. In the case of the Chopin B flat minor sonata, many of these functions and points are not at all obvious from just looking at the score. Consequently, it is vitally important to hear the piece, particularly a number of different performances of the piece so that information derived from identifying points of agreement between different performances can be compared and analysed. A number of different hypotheses as to what these points of change and contrast might be are in the process of being investigated. Possible candidates would be subtle changes in the timing information or in changes in dynamic detail or some combination of the two. Comparison of different performances will be important in order to identify not just the correspondences between points of change in different performances but the extent and limits of the degrees of change.

2 Performance Analysis

Although there is few pieces of traditional musical analysis in existence, there is a wealth of audio recordings of the full sonata. Using performance analysis provides us parameters to measure in a quantifiable way, which can help to establish the performance tradition of the piece. Despite any particular performer's deviations from the score which form their own interpretations of the piece, there are still regularities from which we can establish patterns.

Previous approaches to performance analysis have mainly been concerned with analysing expressive performance patterns to distinguish between different performers or the use of score-tracking, which tracks the performance in terms of the musical score and provides a computer accompaniment for the soloist.

The Performance Worm system created by Langner and Goebl [LG02] and later used by Dixon [Dix03] plots a 2D graph of dynamics versus tempo for each performance that it measures. The dynamics are measured from the spectrum of the audio signal and the pulse is extracted from the same audio signal using the beat-tracking system Beat-root. The musical timing of the notes relative to their expected time and duration can then be calculated. The resultant graphs of dynamics versus tempo allow easy comparison of two different performers playing the same piece,however, it does not provide useful information about the particular performance itself in terms of the musical score.

In Eric Clarke's experiment, [Cla98] measuring a professional pianist playing a Chopin prelude six times on a Yamaha MIDI grand piano, tempo and dynamics are plotted with respect to time i.e. the place in the score. The performer was given no directions to vary his interpretation of the piece or to stick to one interpretation. The six resultant performances therefore differ where the performer has picked out different passages of interest. He says of the expressive performance parameters:"the force of musical expression must be understood by interpreting the function of any expressive features within the specific structural context that they occur. What may appear to be the same expressive element - an acceleration for instance - may have quite opposed functions depending on the structural context in which it occurs" [Cla98]. This demonstrates how structural context is extremely important when considering the different performance parameters across various performances.

Performance research has greatly benefited from the development of score-performance matching, which relates performance notes to score notes in an effort to extract expressive timing patterns [CLB99].However, in most score-matching cases, the information of the structure of the music is not considered.

Dannenberg uses score performance matching to create a synchronized computer accompaniment to accompany solo instruments [DM88]. The system uses MIDI for input and output, so devices such as pitch detectors or keyboard controllers are used to detect the solo performance and the output device are usually synthesizers. Another system using score-performance matching is Curry's co-operative performance system [CW99],which measures the audio information from the performer, calculates the tempo deviations from the score notes as well as the average tempo, follows the performance on the score and attempts to accompany it in time. The matching process in this system is aided by the strict structure of events in music but again does not take great detail from the musical structures that the score represents.

Other information to be measured from a performance can include the field of visual gestures. From the studies of performance gestures of the pianist Glenn Gould, Delalande proposes a three level stucture of gestures ranging from functional to abstract[Del88]. The first level are effective gestures, which are necessary for playing the instrument i.e. bowing, blowing, pressing keys etc. Accompanist gestures are those movements which are associated to effective gestures i.e. elbow and chest movements etc. which are used to help the performer articulate a particular sound. The final level is figurative gestures which are visually perceived by the audience but seem to have no correlation with the actual production of the sound. These gestures are usually captured using motion analysis techniques from video footage of the performance.

3 Acquiring multi modal data

In order to fully establish performance traditions, the maximum amount of data possible must be analysed for each performance. From the wealth of performance parameters used in traditional western classical music, this approach will focus on movement,dynamics and tempo. This data will be captured from audio, video and MIDI recordings made of the various performances of Chopin's B flat minor sonata finale by a collection of professional pianists. The audio, video and MIDI streams will all be acquired through separate applications. Previous analysis has usually focussed on just one of these streams of data and as Eric Clarke has said "there is little work that tries to bring these different sources of information together to provide a more multidimensional view of performance"[Cla04].

MIDI

In order to retrieve MIDI information from the performances, the MOOG piano bar will be used. This consists of a scanner bar that sits slightly above the piano keys of any regular 88 key piano and a pedal sensor that rests underneath the piano's pedals. The scanner bar sits against the fall board of the piano and is designed to be completely out of the way so as not to disturb performers. The scanner bar measures the deflection of the keys in order to sense which keys have been pressed and then to determine their velocities. The sensors feed the note and the velocity of the note to the Control Module where it is transformed into MIDI information. It is then recorded through the open-source sequencer Rosegarden. The MIDI will provide information on what key is pressed, its onset time, its offset time, the key velocity and the time at which either the sostenuto pedal or the una corda pedal is pressed.

Audio

The audio recordings are acquired through the open-source application Ardour and uses Jack for low latency audio I/O which is around a few milliseconds on some computers. This audio data can provide information that is unavailable on the MIDI stream, for example, the timbre of the sound produced.

Video

The video recordings of finger movement across the keyboard of the piano are made on an AVT Guppy F-046C camera. A skilled pianist can play up to 30 sequential notes per second over extended passages[RN82] and as this camera can only capture up to 53 frames/sec at maximum resolution, it is barely sufficient even when supplemented by timing information from the audio and MIDI streams. However, it has a Region of Interest facility(RoI) which permits a smaller amount than the full frame to be sent "flat out". In the case of recording video for the piano performances, as the camera will be recording movement over the horizontal plane of the keyboard, the region of interest can be narrowed to a long, thin frame which can then record in excess of 60 frames/sec. The performer's more general body gestures such as elbow or upper torso movement is then captured by another camera at maximum resolution. The video information is then used to analyse the performer's movements i.e. the musical gestures that take place in the performance. This can be achieved with traditional motion capture techniques tracking markers on the performer. These markers, which in this case will be blobs of colour, will be placed on the joints of interest i.e. knuckles of the fingers, wrists and elbows to determine their curvature and also to track their movement without causing disturbance to the performance.

4 Analysing multi modal performance data

The performances are analysed from the multi-variate data for the various performances,that is the audio, video and MIDI streams. Performance notes are extracted from the recorded performance to compare with the score notes in order to allow the computer to track where the performance is, and calculate expressive timing deviations. Since key velocity correlates directly with the loudness of the sound being produced, the dynamics of the piece are extracted from the key velocity information captured by MIDI and the loudness curve produced can be validated by the audio recording. Parts of the piece that have been highlighted through performing accented notes or the illumination of a secondary melody can be seen from the MIDI information in large differences of neighbouring note's velocities. Musical performance gestures that are captured on video can be analysed in the three categories as defined by Delalande earlier in the paper. Detailed analysis of expressive performance parameters have previously neglected to consider the integrated effect and the interaction of performance parameters on one another and so we must be careful not to consider each parameter only separately.

This tempo, dynamics and gesture information allows us to look at the segmentation of the piece in terms of musical analysis and compare it to the traditional analysis available. Due to the unbroken rhythm in this piece, tempo analysis will be easier to perform as nearly every note is described in the score as an equal measure. Any deviation can therefore be considered to be illuminating a part of the piece for reasons of structure or otherwise. For example, when using tempo to convey the structural issues of the composition, a performer may slow down before a boundary e.g. before the entry of a second theme. In the passage of music printed below, audio recordings such as that of the pianist Lief Ove Andsnes indicate that the performer is slowing the tempo slightly and lowering the dynamic level before entering the new theme in bar 23. Other audio recordings such as the one by the pianist Ignaz Friedman indicate the performer has made a crescendo into the new theme but emphasises the A flats in bars 23-24 by holding back the tempo slightly on these particular notes.

5 Data Storage

Storing the data in a way which is amenable to future extraction is a far from trivial process, especially as it would be highly beneficial to be able to compare a range of different performers against the benchmark positions indicated in the score. Many data formats have been proposed which can represent musical score, MusicXML[Goo01] being among the most popular. Storing precise timing information in tandem with audio is also relatively straight forward; the simple open-source audio editor Audacity[Aud] supports audio file tagging, for example, and can read and write tags to text files. To integrate all of this information and to provide an infrastructure for more sophisticated subsequent data processing, the choice is more limited: current solutions include the Music Encoding Initiative (MEI)[Per02] and Performance Mark-up Language (PML).

MEI's main intention is to "a) provide a standardised universal XML encoding format for music content (and its accompanying metadata) and b) facilitate interchange of the encoded data". It can represent a score; it has some facility for the representation of logical and analytical domain information; but there is facility for time-stamping objects in various time codes. However, the semantics assoicated with time-stamp objects are obtuse, and there is no explicit separate representation for performance data.

Conversely, PML was particularly designed to accommodate the mark-up of performance information. PML is a spefcification which can be used to extend XML-based score representations such as Music-XML. Analytical, performance and score information are separated into different hierarchies. Since MEI represents these domains in a single hierarchy which is based on the requirements of the features of the musical score, it makes it a less elegant solution for the representation of other data which may be non-isomorphic with the score. For example, one would not expect the repeated portion of a da capo aria to be performed the same way the second time.

6 Conclusion

After acquiring, analysing and storing the performance data, comparisons between the separate performances will provide us with the performance traditions for the piece, which can then reveal the more subtle structural issues that have previously been difficult to extract from simply looking at the score. Conclusions can then be drawn from the comparison of the multi-dimensional performance analysis to the more traditional musicological analysis.

References

[Aud] Audacity. Free software foundation : Gnu public license. http://audacity.sourceforge.nets.

[Cla98] Eric F. Clarke. The semiotics of expression in musical performance. *Contemporary Music Review*, 17, Part 2:87–102, 1998.

[Cla04] Eric Clarke. *Empirical Musicology : Aims, Methods, Prospects*, chapter Empirical Methods in the Study of Performance, pages 77–102. Oxford University Press, 2004.

[CLB99] P. Cano, A. Loscos, and J. Bonada. Score performance matching using hmms. *Proceedings of the ICMC*, 1999.

[CW99] Ben Curry and Geraint Wiggins. A new approach to cooperative performance :a preliminary experiment. *International Journal of Computing Anticipatory Systems*, 4:163–178, 1999.

[Del88] Francois Delalande. *Glenn Gould,Pluriel*, chapter La gestique de Gould:elements pour un semiolgie du geste musical, pages 83–111. Louise Corteau Editrice Inc, 1988.

[Dix03] Simon Dixon. On the analysis of musical expression in audio signals. *Storage and Retrieval for Media Databases, SPIE-IS&T Electronic Imaging*, 5021:122–132, 2003.

[DM88] R Dannenberg and H Mukaino. New techniques for enhanced quality of computer accompaniment. *Proceedings of the International Computer Music Conference*, pages 243–249, 1988.

[Goo01] Michael Good. *The Virtual Score:Representation, Retreival, Restoration: Vol.9.*, chapter MusicXML for Notation and Analysis, pages 113–124. MIT Press, 2001.

[LG02] Jorg Langner and Werner Goebl. Representing expressive performance in tempo-loudness space. *ESCOM Creativity, Liege, April 2002*, 2002.

[Per02] Roland Perry. The music encoding initiative (mei). *Proceedings of the First International Conference on Musical Applications Using XML*, pages p55–59, 2002.

[RN82] D.E Rumelhart and D.A Norman. Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, 6:1–36, 1982.

[Ros99] Charles Rosen. *The Romantic Generation.* Fontana Press, 1999.

[Tal01] Michael Talbot. *The Finale in Western Instrumental Music.* Oxford University Press, 2001.

[Tan05] Richard Taniskin. *The Oxford History of Western Music, Vol3:The 19th Century.* Oxford University Press, 2005.