VISUALISING MUSICAL STRUCTURE THROUGH PERFORMANCE GESTURE

J.MacRitchie, B.Buck and N.J.Bailey
Centre for Music Technology, University of Glasgow
{j.macritchie, b.buck, n.j.bailey}@elec.gla.ac.uk

ABSTRACT

A musical performance is seen as the performer's interpretation of a musical score, illuminating the interaction between the musical structure and implied emotive character [1]. It has been demonstrated that performers' physical gestures correlate with structural and emotional aspects of the piece they are performing and that this information can be decoded by an audience when presented with a visual-only performance [2].

This paper investigates the relationship between direction of physical movement and underlying musical structures. The Vicon motion capture system is used to record 3D movements made by nine university-level pianists performing Chopin preludes op.28 Nos 6 and 7. The examination of several pianists provides insight into the similarity and differences in gestures between performers and how these relate to structure.

Principal Component Analysis (PCA) of these performances and consequent analysis of variance reveals a relationship between extrema of the first six significant components and timing of phrasing structure in Prelude 7 where motion troughs consistently lag behind the occurence of phrase boundaries in the audio. This relationship is then examined for Prelude 6 which encompasses longer, expanded phrases and changes in rhythm. These expanded phrases are associated with elongated or split gestures, and variations of the motif with changes in movement.

1. INTRODUCTION

Structural communication in performance is well understood for parameters such as timing and dynamics, and certain relationships between these parameters and structures have been clarified [3]. We now know that performers tend to slow the tempo towards the end of a phrase and use dynamics to 'shape' a phrase often using a diminuendo towards the end. These are of course context-dependent as a performer can use the same parameters to mark different structural features [4]. Performers also have personal styles of playing and will not all use the same performance parameter to emphasise the same structural feature.

This personal style becomes a much bigger issue when treading into the field of physical gestures. No such straightforward relationships exist between physical gestures and musical structure, although it has been demonstrated that performers' gestures do contain information about the music being performed [5, 6]. Perception studies have also shown that typical audience members can accurately perceive information about tension and phrasing from visual-only performances [2].

This paper aims to establish relationships between body movement and the phrasing structure it attempts to convey, exploring how these change between different pieces of music. This investigation of phrasing and gesture relationships will be conducted by recording piano performances through the Vicon motion capture system in synchrony with audio recordings, and subsequent analysis of the movement results alongside traditional analyses of structure in the chosen pieces.

2. CHOPIN PRELUDES

Several factors fuelled the choice of two Chopin preludes for gestural analysis.

- In order to make general statements about phrasing structure, it was necessary to provide some amount of scientific control, otherwise the exercise could become tantamount to guesswork. Two pieces were therefore chosen to examine the progression of movement and structure relationships.
- The pieces' genre may have an effect on performance gestures, so ideally those from the romantic period would provide the most expressive performances.
- The motion capture system coped better with shorter recordings and so brief pieces were preferred.

The preludes chosen for this investigation were Preludes op.28 No.s 6 & 7. Prelude No. 7 in A major has a strict, rigid structure, with the rhythmically identical two bar phrase occurring eight times in total. As can be seen from Figure 1, this binary form 16-bar piece has the main boundary occurring exactly halfway through at bar 8, and a harmonic arrival point occurring with the chord at the end of bar twelve [7]. The two sections of the piece are thought to each contain a set of antecedent-consequent phrases.

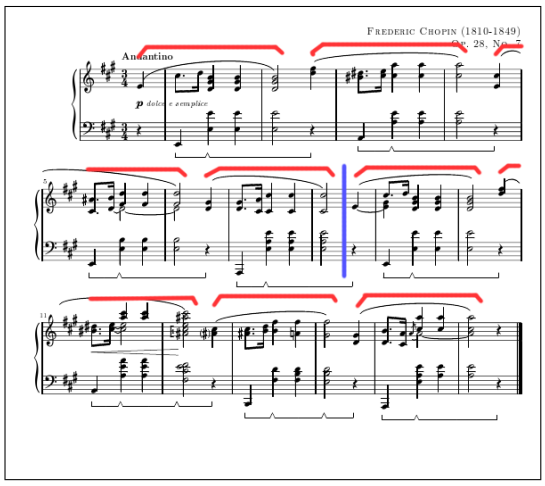

Figure 1. Chopin prelude op.28 No.7 with two-bar phrases marked in red and line boundary for end of first section.

Prelude No.6 in B minor comprises three sections [8] from bars 1-8, bars 9-22 and a coda section from bars 23-26. The first section represents an 'extended idea'. As seen in Figure 2, Chopin begins with another two-bar motif in B minor. This motif is repeated with a slightly higher pitch range in the next two bars. The first part of the motif is repeated again and then expands into a four bar phrase ending at bar 8, the first sectional boundary. The second section represents an expansion of this idea. At bar 9, The two-bar motif from the beginning is repeated with the next expansion moving into C major. A new four bar phrase is introduced at bar 15, answered by the consequent four bar phrase concluding on the tonic at bar 22 at the second sectional boundary. The piece ends with a slight coda in B minor in its final phrase [1]

As many different analyses of one piece can exist, each pianist's own interpretation of phrasing is noted within this experiment. Analysis of each pianist's performance of the rigidly structured prelude no.7 will provide an impression of performance style. This piece provides the opportunity to observe movements for each phrase in isolation before moving on to examine the other prelude containing slightly more complicated structures.

3. METHOD

3.1 Performance Motion Capture

Using the Vicon 3D-motion capture system [9], performances of the two selected Chopin preludes by nine highly trained pianists [2] were recorded. All pianists were asked

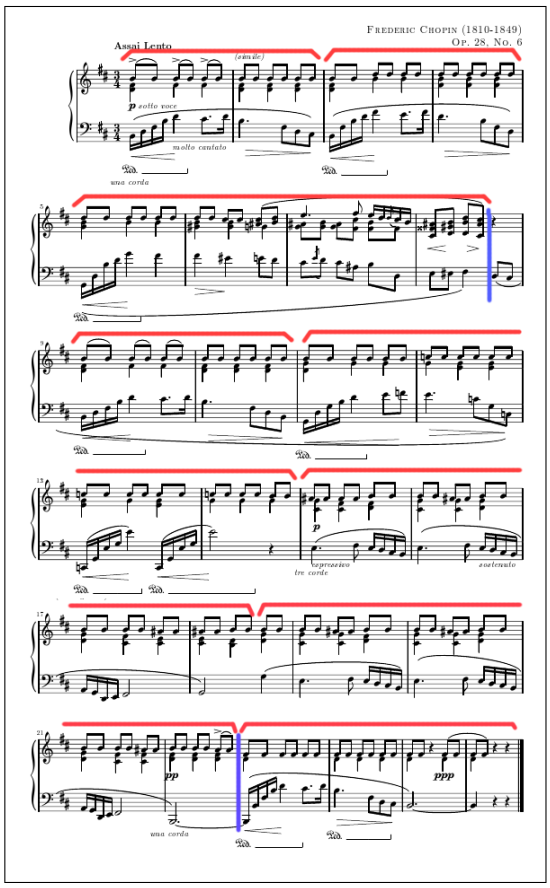

Figure 2. Chopin prelude op.28 No.6 with phrases marked in red and line boundaries for the ends of sections.

to play the pieces from memory in an effort to ensure an in-depth knowledge of both pieces. The only performance direction given to the pianists was to play as if they were in a normal concert setting. Pianists' interpretation of phrasing structure and gestural expression were taken by means of a self-report following the recording.

The Vicon motion capture system consists of twelve infra-red cameras placed around the room to ensure capture of a particular volume of space. Retro-reflective markers were placed onto a velcro jacket and hat worn by the performers in the configuration shown in the head and upper body model in Figure 3. This particular model combined the upper body model from Cutti et al. [10] with four reference markers for the head positions. Each camera tracks the coordinates of the 28 markers and triangulates their position in order to build a 3-D model of each performer. Each video was recorded at 120 frames per second in synchrony with an analogue input for the recorded sound. The models were then reconstructed by post-processing and any points where the cameras had failed to pick up a certain marker were filled with the estimation models available from the Vicon Nexus software.

Problems were encountered particularly with the markers placed on the elbows of the performers. As the markers were placed not directly onto the skin but onto a velcro jacket, there were several points in the recordings where the marker was lost by the camera as the jacket had moved

[1] This analysis of Chopin's Prelude Op.28 No.6 is combined from Kofi Agawu, V. 'Concepts of Closure and Chopin's opus 28' in Music Theory Spectrum 9:1–17, 1987. and comments made by Jennifer MacRitchie, University of Glasgow, and David Lewis and Christophe Rhodes, Goldsmiths, University of London

[2] These nine performers consisted of five music performance undergraduate students, four at the University of Glasgow and one at the University of Edinburgh, two postgraduate students from the Royal Scottish Academy of Music and Drama and two amateur pianists with more than ten years of performance experience. Each pianist was paid a one-off sum of £25 for their participation in the experiment.

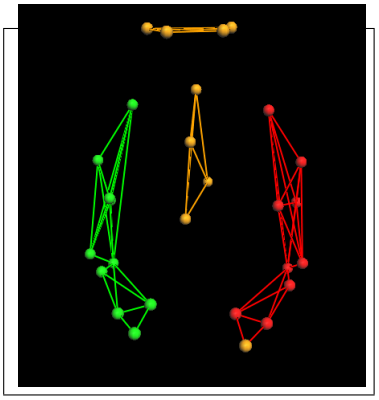

Figure 3. Vicon marker model for upper body of pianists.

round the elbow and displaced the marker. Although the Vicon interpolation algorithms filled most of these elbow gaps, the system is proprietary, so these algorithms are unavailable for inspection. The accuracy of reconstruction for these elbow gaps must therefore be considered suspect.

4. RESULTS

4.1 Motion Capture Analysis

As motion capture always produces such an overwhelming plethora of data, the traditional phrase analysis of each prelude provides us with points from which to start investigation of gestural cues at phrasing boundaries. Each performer's audio recording was annotated in Audacity [11] with the timings of the phrase boundaries explained in section 2. by a separate professional pianist. Each performer's own view of the phrase segmentation was also noted in case of any differences to traditional analysis. The pianists' self-reports also conveyed a wide view on the role of movement in performance, with some branding movements extra to sound productive ones as completely unnecessary and something they tried to limit, whilst others felt it vital to move in order to 'feel' the music they were performing. Although physical gestures in performance can be classified as movements necessary to the actual sound-production, or movements that are related to the music but not necessary for the actual sound (i.e. ancillary) [12], it is acknowledged that gestures may still be multi-functional. To view the overall general motion characteristics of each performer, principal component analysis (PCA) was performed through designated pca modules using singular value decomposition algorithms [13] on the complete set of motion data for each pianist. Each person's principal component score was mapped against the timings of each phrasing boundary to determine if there was a pattern of movement for each phrase. Reduced-dimension curves such as these are good at expressing a general overview but tend to lose some semantics of the actual movement being performed and so each individual marker is then also examined for reference to phrases, measures and beats.

Three pianists have been chosen to demonstrate the spread of results concisely. These pianists were chosen according to their ability, their standard deviation and variance of movement calculated for intra-performance data and also their views on movement during a performance. Performer 1 is a highly trained amateur pianist and had a small standard deviation of movement. Performer 2 is a conservatory trained postgraduate student and had a large standard deviation of movement, and Performer 3 is a music undergraduate student and had a mid-range standard deviation. Normalization of results allows the movements to be correlated with phrase structure independent of differences in amplitude. The arrows in each graph indicate the point in time where the last note of each phrase ends in the audio stream.

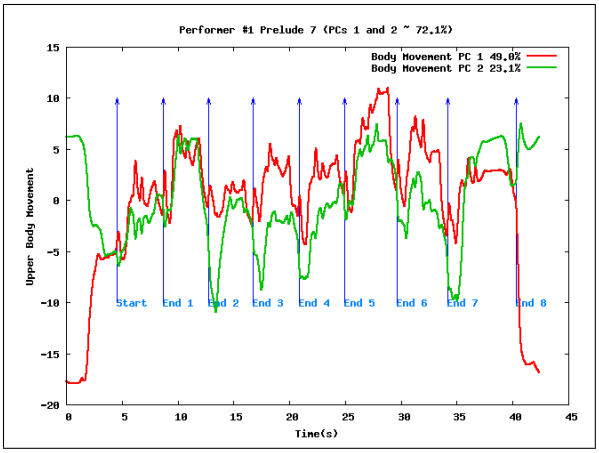

Figure 4. First two principal components of movement for prelude 7, Performer 1 mapped against phrase boundaries.

4.2 Prelude 7 in A major

Starting with prelude 7, in which the pianists self-reporting analysis agreed with the traditional phrase segmentation marked in section 2., Figures 4, 5 and 6 show the first two principal components accounting for around 70% of the overall movement. These appear to relate to the phrase boundaries as dictated by traditional music analysis. For each performer, the correlation between markers and the resultant PCA curves i.e. the loadings, clarified that instead of a few markers being prevalent in causing the most variance in motion, the PCA curves were a result of the variances in a combination of several markers and these differed slightly for each pianist.

Interestingly, Performer 1's self-report on conclusion of the recordings expressed the opinion that movement in performance did not convey any information on phrasing and that during performances, he/she attempted to minimize movements and facial expressions. However, Performer 1's movement, shown in Figure 4, shows a clear relationship between physical gestures and phrasing structure where each phrase boundary precedes a trough in the motion graph. The loadings for Performer 1 related highly to movements in the upper arms, the elbows, the wrists and the chest.

Performer 2's main component loadings consist of movement in every section of markers: the wrists, elbows, upper arms, shoulders, chest and the head. A pattern can

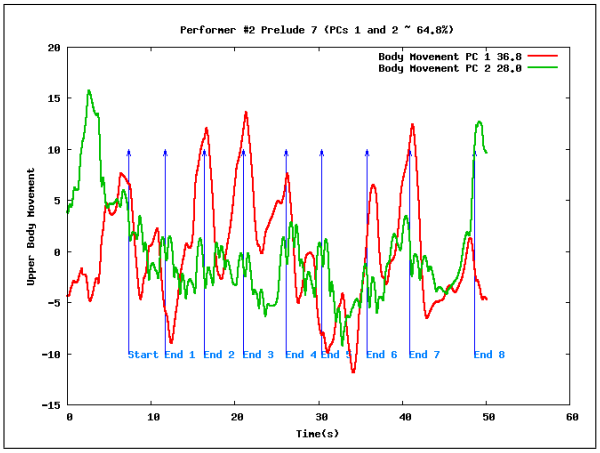

Figure 5. First two principal components of movement for prelude 7, Performer 2 mapped against phrase boundaries.

be seen in Figure 5 for the first principal component(red), where the phrase boundary occurs at a motion peak for all but phrase endings 1, 5 and 6 whose peaks precede the boundary in time. The second component (green) shows a relationship between the phrase endings and peaks in the graph.

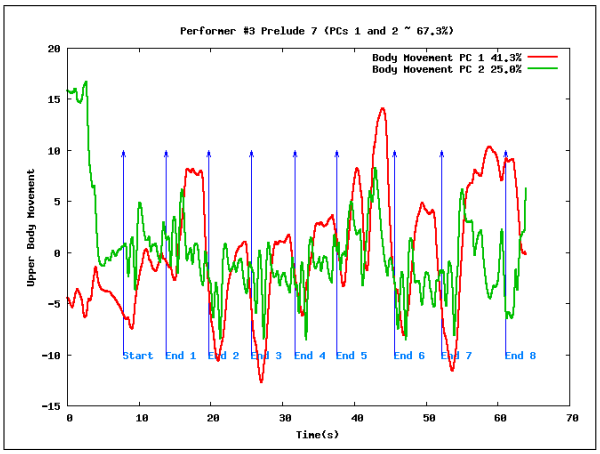

Figure 6. First two principal components of movement for prelude 7, Performer 3 mapped against phrase boundaries.

Performer 3's main component loadings relate to movements in the wrists, upper arms, chest and head. A pattern can be seen in Figure 6 where each phrase boundary precedes a trough in the main component in all phrases. The pattern of movement for each phrase in this component is changed slightly during phrase six. This could be to emphasise the harmonic arrival in bar 12 of the piece.

The addition of the weighted values of the first six principal component scores for each performer produces a visualisation of data accounting for more than 90% of the variance in motion, the weightings calculated from the percentage variance of each component over the dataset. These have been resampled with 10,000 points and time-warped to allow more direct comparison between performers. The distance between each audio phrase boundary is 0.1 and quoted means and standard deviations are calculated for the distances between the troughs of the motion trajectory and its corresponding phrase boundary. Figure 8 shows a pattern in all phrases except one, five and six (mean=0.0369,s.d=0.015). Phrases 1 and 5 are the first phrases in the each section of prelude 7 whilst phrase 6 contains the harmonic arrival point. Figures 7 and 9 show a clear relationship between the movement and phrase boundaries.(mean=0.0186,s.d=0.0116 and mean=0.0204,s.d=0.0068 respectively) The calibration of this hypothesis with prelude 7 now provides us with a useful tool to observe the structure of prelude 6 for the same performers.

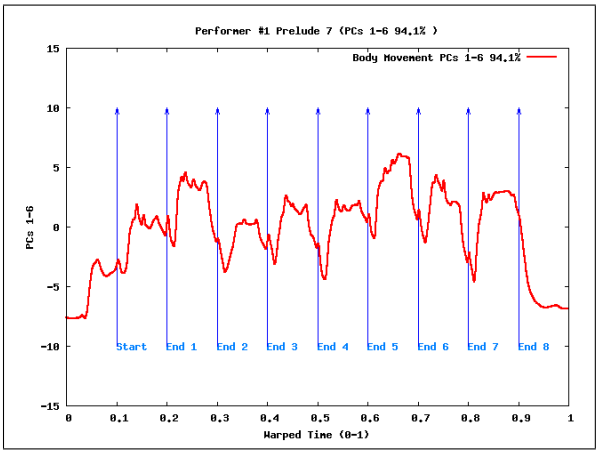

Figure 7. Combination of first six components for prelude 7, Performer 1 mapped against phrase boundaries.

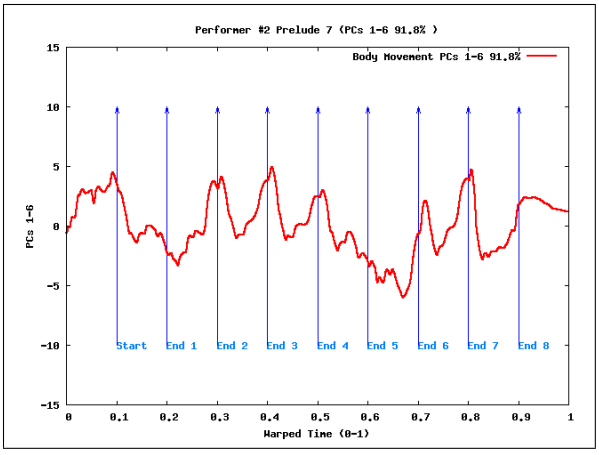

Figure 8. Combination of first six principal components of movement for prelude 7, Performer 2 mapped against phrase boundaries.

4.3 Prelude 6 in B minor

The initial two-bar motif in prelude 6 is in the left hand melody marked in the score seen in section 2. This motif is varied in the subsequent phrases, first in pitch for the second phrase, then also in rhythm for the third phrase ending at bar 8. This is echoed in the movements made by performers. This initial step looks at the first three

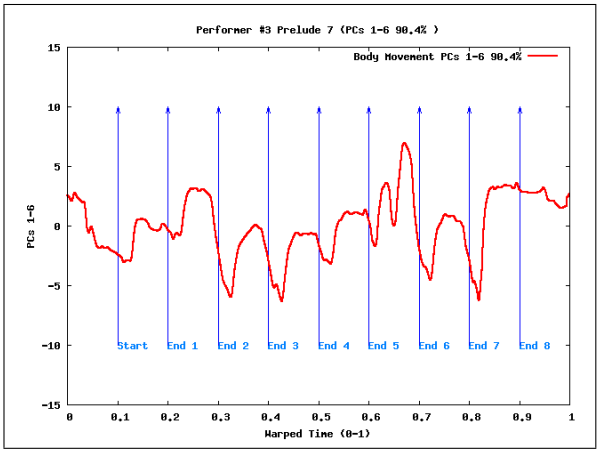

Figure 9. Combination of first six principal components of movement for prelude 7, Performer 3 mapped against phrase boundaries.

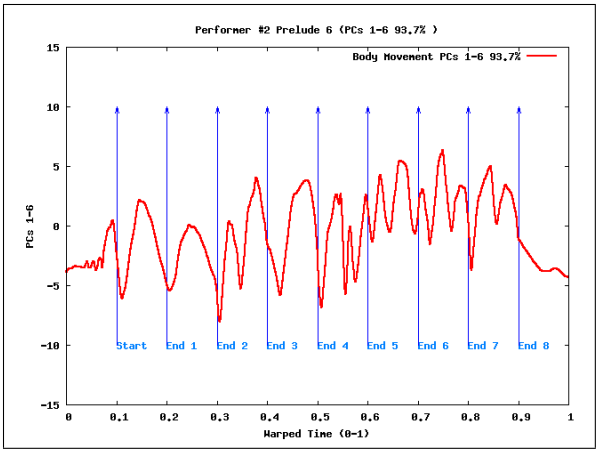

Figure 11. Combination of first six principal components of movement for prelude 6, Performer 2 mapped against phrase boundaries.

phrases of prelude 6 as agreement on phrase segmentation between analyses and performers' self-reports diverge from this point onwards. The means and standard deviations of distance between motion trough and phrase boundary are for the first three phrases only.

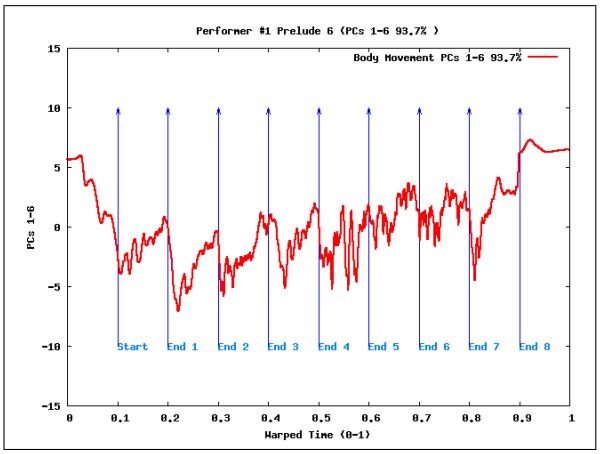

Figure 10. Combination of first six principal components of movement for prelude 6, Performer 1 mapped against phrase boundaries.

Performer 1's weighted combination of principal components of movement, shown in Figure 10, shows a distinct pattern of movement for the two-bar motif established in phrase one of the piece (mean=0.0223,s.d=0.0110). Its elongation in phrase 3 is mimicked by an elongated gestural movement. The change in motif at bar 15 beginning with a four bar phrase (phrase 5), is marked with a different pattern of movement. This movement is repeated as the consequent four bar phrase is played.

The first six principal components for Performer 2, as seen in Figure 11, also shows a clear pattern within the first three phrases (mean=0.0129,s.d=0.0115 this particular mean is negative as each trough occurs slightly before the phrase boundary). Interestingly, in phrase 3 where the original two-bar motif is expanded, we clearly see two separate movements. As the length of the phrase being performed is just under 12 seconds long, we refer to the theory of gestures being separated into gesture-units i.e. action-chunking [14]. At which points within a long phrase this action-chunking occurs is most likely related to the smaller rhythmical groupings within the particular phrase.

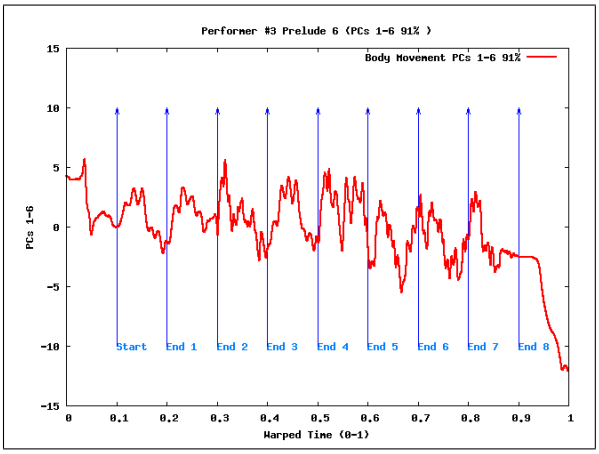

Figure 12. Combination of first six principal components of movement for prelude 6, Performer 3 mapped against phrase boundaries.

The first principal component for Performer 3 as seen in Figure 12 again shows this pattern between movement and phrasing (mean=-0.0069,s.d=0.0086). Again at phrase 3, we can see the beginning of the two-bar motif related movement before the phrase is expanded and the resulting gesture elongated as well.

4.4 Analysis Across All Performers

One-way ANOVAs were conducted to investigate the effects of performer style and phrasing motion. No significant effect of performer was found which, when coupled with variance analyses of individual performers, revealed that performers were consistent in their movement

patterns throughout the piece, showing no significant difference between performers in their overall motion timing variance. A significant effect of trough location was found between performers ($F=17.32$, $p<0.001$). When considered with the minimal variance within performances, these results show that performers were consistent in the location of their coherant motion patterns (troughs) with respect to phrasing boundaries. As performers' interpretations of the latter part of prelude 6 differ, such straightforward comparisons cannot be performed for this piece.

Comparisons of pianists' singular marker movements for both preludes showed similarities in movement in the y axis (along the length of the keyboard) of every single marker of the upper body. Troughs in each plot occurred either slightly before or slightly after the audio phrase boundary. This changed between pianists but was consistent between performances of two preludes with differing structure and different melodies. Differences in pianists' marker movements showed for some pianists, a clear relationship with phrasing in all three axes of movement for each marker. Some used their heads to mark out phrasing whereas others preferred to use their upper body. These movements were performer specific and occurred across both pieces. Within some markers denoting phrasing, movements corresponding to the measures and beats within the piece were found. Markers which were not phrasing specific also appeared to highlight beats or measures of the piece. Despite this, the overall general movement reflected by the PCA shows a clear phrasing pattern.

5. CONCLUSIONS

Structural information appears to be inherent in pianists' directional movements across the three axes. Principal component analysis confirms the relationship between general movement in the upper body and head with composed structure. Variance analysis shows that each performer's general movement consistently lags behind the occurence of a phrase boundary in the audio stream. By examining pianists' movements in performances of Chopin's prelude op.28 no.7, it is confirmed that short phrases in isolation, with the same rhythmical pattern appear to invoke similar movements by the performers. Movement also appears to change when the motif is varied, as examined in performances of Chopin's prelude op.28 no.6. In comparison with the short two-bar motif, phrases with a longer duration have different elongated gestures and are sometimes split into sections in a process referred to as action-chunking.

This investigation provides the initial step of relating movements to phrases. Further investigations into effect of genre of the music being performed on the structural relationship with movement are required to make more general statements across a wider variety of music. Further steps for this research are to clarify whether physical gestures are related to other compositional or performance attributes. Empirically relating general movement to structural aspects of performed music contributes to the argument that ancillary performer movements may have a music-related function. This has implications for piano pedagogy and furthers the understanding of the wider relationship between music and movement.

6. ACKNOWLEDGEMENTS

The authors acknowledge Prof. Frank Pollick and researcher Lukasz Piwek of the Psychology Dept at the University of Glasgow for their support and access to equipment. This research is EPSRC funded.

7. REFERENCES

[1] Shaffer, L. Musical Performance as Interpretation *Psychology of Music*,23, 17–28, 1995.

[2] Vines,B. et al. Cross-modal Interactions in the Perception of Musical Performance. *Cognition*, 101, 80–113, 2006.

[3] Friberg, A. and Battel, G.U. Structural Communication in Parncutt, R. and McPherson, G.E. (Eds.) *The Science and Psychology of Music Performance*, Oxford University Press, 2002, 199–218

[4] Clarke, E.F. The Semiotics of Expression in Musical Performance *Contemporary Music Review*, 1998, 17, Part 2, 87–102

[5] Wanderley, M.M. Quantitative Analysis of Non-Obvious Performer Gestures Gesture and Sign Language in *Human-Computer Interaction: International Gesture Workshop*, 2001.

[6] Jane W. Davidson. Qualitative insights into the use of expressive body movement in solo piano performance: a case study approach. *Psychology of Music*, 35(3):381–401, 2007.

[7] Kresky, J. A Reader's Guide to the Chopin Preludes Greenwood Press, 1994.

[8] Kofi Agawu, V. Concepts of Closure and Chopin's opus 28 *Music Theory Spectrum*, 9:1–17, 1987.

[9] Motion Capture Systems from Vicon `http://www.vicon.com`.

[10] Cutti,A.G. et al. Soft tissue artefact assessment in humeral axial rotation *Gait and Posture*, 2005, 21, 341-349.

[11] Ardour - the new digital audio workstation. `http://www.ardour.org`.

[12] Cadoz, C. and Wanderley, M. Gesture-Music in Wanderley, M. and Battier, M. (ed.) *Trends in Gestural Control of Music* Iracam - Centre Pompidou, 2000

[13] Risvik, H. PCA Python module `http://folk.uio.no/henninri/pca_module`

[14] Godøy, R. I. Knowledge in Music Theory by Shapes of Musical Objects and Sound-Producing Actions. In M. Leman (ed.) *Music, Gestalt, and Computing*, Springer Verlag, Berlin (1997) pp. 106–110