MOTION RECURRENCE ANALYSIS IN MUSIC PERFORMANCES

Euler Teixeira, Hani Yehia, Mauricio Loureiro
Universidade Federal de Minas Gerais, CEFALA
Av. Antonio Carlos 6627, Belo Horizonte, Brazil
euler.teixeira@cpdee.ufmg.br

Marcelo Wanderley
McGill University, IDMIL - CIRMMT
555 Sherbrooke St. W., Montreal, Canada
marcelo.wanderley@mcgill.ca

ABSTRACT

This work presents a method to represent, segment and analyze the recurrence patterns on motion data during musical performances. Physical gestures were extracted during clarinet performances and analyzed according to gestural features, comparing different musicians, musical passages and performance styles. The gestural aspects of the performances were related to the musical structure and its expressive content, and an acoustical analysis validated the results. Results show a recurrent sequence of clarinet gestures inside a defined region of interest, shown to be a key moment in the music.

1. INTRODUCTION

Musical expressiveness is a concept that is hard to be formalized by objective data. There has been recently a growing search for methods to describe and analyze it according to a set of quantitative parameters. This has been done mainly through the audio analysis of musical performances, extracting musical content information directly from the acoustical data [1, 2]. Studies have shown that musicians make use of small deviations, regarding note durations, articulations, intensity, pitch and timbre, in order to convey their musical intentions [3, 4].

This study expands acoustical analysis methods for investigating musicians' expressive intentions, incorporating information about their body movements during musical performances. We present a method to define and analyze the physical gestures executed by the musicians while playing their instruments, and to extract motion parameters that can be objectively related to their expressive intentions and to the musical structure [5, 6]. This sort of multi-modal investigation has also been successfully employed in studies related to the analysis of speech [7] and dance [8], examining the coupling between their acoustical and visual components.

Observing a musical performance, it is possible to notice that the body movements executed by the musicians, besides being in many cases essential to the instrument's sound production itself, are also closely related to the musician's expressive intentions in a particular performance [9, 10]. Even with some recent studies in this direction [11–14], there is not so far a unique and objective method that can be widely used to extract and analyze such information from the motion capture data. Despite this, there is strong evidence that such expressive information is present in musicians' body movements, providing valuable information to better comprehend expressiveness from a multimodal point of view [15, 16].

In order to establish relations between performers' body movements and their expressive musical intentions, there are three key steps [6, 17]. The first step is to track points of interest in the musicians' body and instrument, during several musical performances, searching for patterns of temporal and spatial evolution, in order to define significant and recurrent physical gestures. After that, it is necessary to compare the gestures of different performers, performance styles and musical passages, taking into account spatial, temporal and musical parameters. The final step is to conduct an analysis over the corresponding acoustical data, searching for related parametrical patterns coupled with the motion analysis.

With this method, it is possible to investigate where the expressive content can be found in the musicians' body movements, what is its behaviour, and how it relates to the musical structure, ultimately defining a musical meaning for the physical gestures of musicians during performances. In this paper, the proposed method is applied to clarinet players' body movements, performing solo pieces of the classical repertoire.

2. EXPERIMENTAL METHODOLOGY

The objective of the study is to analyze the expressive content of musical performances by a group of clarinet players, based on the corresponding motion and acoustical data. In the proposed experiment, the motion tracking is done with a high-end 3D motion capture device, the Optotrak Certus. It consists of a tracker, built with three infra-red cameras positioned along an axis, giving the spatial position of active infra-red LED markers, inside a tridimensional measurement volume, also permitting a synchronous audio recording.

The studied group of musicians consists of 10 classical clarinet players, 8 males and 2 females. The selected musical excerpt is presented in Figure 1. The musicians were asked to perform it according to two distinct performance styles. First, in the standard style, each musician performed the music freely. After that, they were asked to follow a metronome, set to a tempo estimated from the previous standard performances. The goal was to obtain an ob-

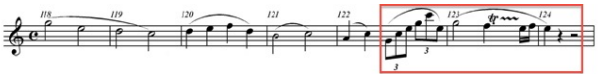

Figure 1. First movement of Mozart's Quintet for Clarinet and Strings in A Major, Kv 581. Final part highlighted.

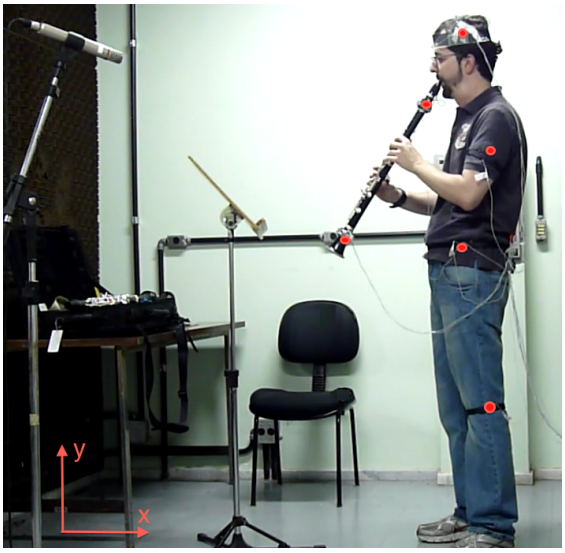

Figure 2. Photograph of the data acquisition experiment, from the motion capture sensor viewpoint.

jective way to compare between free and expressive musical performances, and control performances with lesser expressive content, restrained by the metronome beat.

Each of the 10 clarinet players performed the excerpt 6 times, 3 according to each of the 2 performance styles. They played standing up, facing the Optotrak tracker sideways, and motion capture markers were placed on their bodies and clarinets, according to Figure 2. The motion capture tracker was placed vertically, about 2 meters away from the musician, using a sampling rate of 100 frames per second. The sound was recorded synchronously to digital audio files, through a condenser microphone positioned about one meter away from the clarinet.

The audio of each performance was processed to extract its pitch and energy envelope curves. Note onsets and offsets were detected using a previously developed system [1], in order to visualize the evolution of each performance, according to the musical structure and the individual aspects of each execution.

3. MOVEMENT REPRESENTATION AND SEGMENTATION

The body movement analysis in this study will be based on the clarinet bell motion. The clarinet motion has been the object of previous studies [9, 12] and it is believed to be an important inflection and expressive tool used by expert clarinet players. In order to analyze the clarinet bell tridimensional motion evolution in conjunction with the acoustical data, we need a strong scalar representation of the motion data in time. A simple way to do that is to use the tangential velocity of the clarinet bell marker's trajectory, calculated in this case using the Euclidian distance between the positions of the marker in two subsequent samples. This unidimensional parameter captures a large amount of information from the musician's movements.

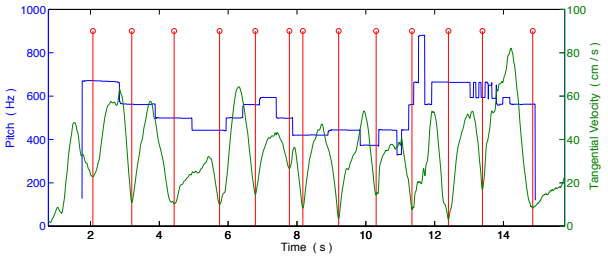

Figure 3. Movement segmentation for a clarinet performance. The blue curve shows the pitch (Hz), the green curve shows the tangential velocity of the clarinet bell (cm/s) and the red lines mark the movement segments.

Through the extraction of pitch and energy envelope curves from the audio signal, it is possible to determine all note onsets and offsets and thus to segment the acoustical data into musical notes and phrases. It is also very important to develop a procedure to segment the movement data accordingly. Unlike the acoustical data, there are no basic units established to segment the movements into, but it is possible to divide those movements into representative segments, according to their geometrical and temporal attributes.

This can be done based on the tangential velocity representation of movement, shown in Figure 3 after appropriate filtering, assuming that its local minima corresponds to inflection points in the musician's movement, where the motion direction or character is most likely to suffer a sudden change [11]. These points were thus used as a basis for the segmentation procedure, defining movement segments between subsequent local minima in the tangential velocity curve.

4. MOVEMENT RECURRENCE

In order to relate the musicians' movements to their expressive intentions and to the musical structure, we analyzed the recurrence of movements within the excerpt, along different performances by the same musician. The search for recurrent patterns in the movements of the musicians was made by using an instantaneous correlation algorithm, developed in [18]. It calculates the correlation coefficient between a pair of signals for each instant in time and also for different time offsets between them, generating a bidimentional correlation map between the two signals, as shown in the bottom half of Figure 4. The horizontal axis represents time, from the first note onset to the last note offset of the music, and the vertical axis represents the time offset between the two analyzed signals, from -0.5 to +0.5 seconds. The blue areas on the map indicate low correlation values between the signals, while the red areas indicate high correlation values.

For each musician, a correlation map was calculated for each of the 15 possible signal pairs of the 6 clarinet tan-

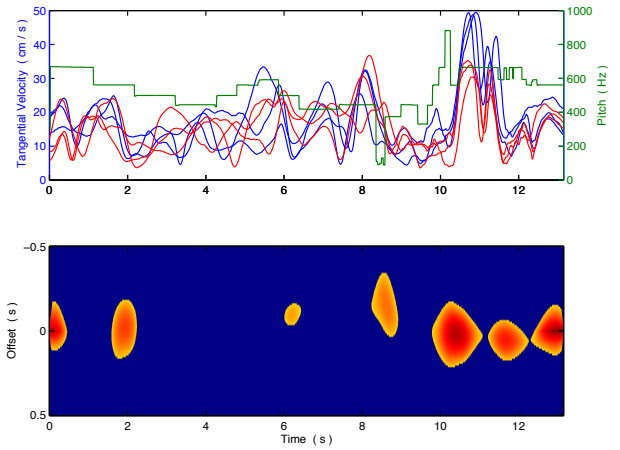

Figure 4. Clarinet bell motion recurrence map for the 6 performances by Musician 1. Top: tangential velocity curves, standard performances in blue, metronome performances in red. Bottom: motion recurrence map, blue indicates low recurrence, red indicates high recurrence.

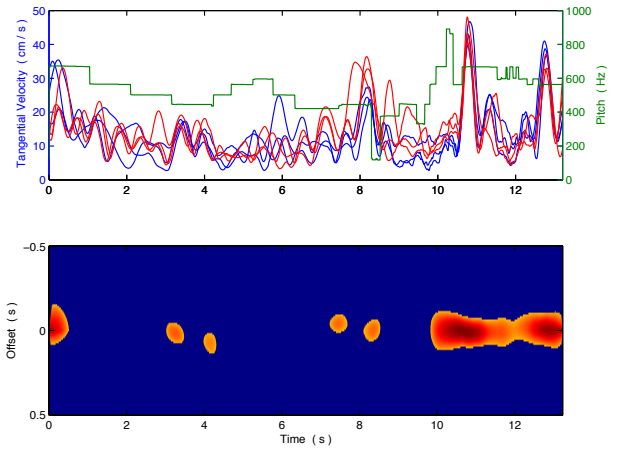

Figure 5. Clarinet bell motion recurrence map for the 6 performances by Musician 9.

gential velocity curves. These 15 correlation maps had their negative correlation values truncated to 0 and were then summed and normalized to 1, generating a resulting map that provides a recurrence measure for that musician's clarinet bell movement over his/her 6 performances. In order to highlight the regions of interest, of high recurrence, a threshold was applied to the recurrence map, removing values below 0.7. Also, to guarantee a perfect temporal alignment between the signals, in accordance with the musical structure of the excerpt, the 6 velocity curves were time-warped [19], using the note onsets as reference points in the timing model. Figure 4 illustrates the result of this recurrence map analysis for Musician 1.

Analyzing Figure 4 it is possible to see the regions in the music where this musician employs recurrent movements along his performances. Most noticeably at the final part, after the 10 seconds mark, where all the 6 velocity curves at the top plot are all highly correlated. This is confirmed by the large dark red areas in the corresponding region of the recurrence map. Some moderate recurrence can also be identified in other regions, but for most of the excerpt, the blue areas on the map indicate the absence of recurrent movement patterns.

This motion recurrence map analysis was conducted for all 10 clarinet players. Figure 5 shows another example. Six of these musicians exhibited high movement recurrence regions in the final part of the music, similarly to the examples shown, two exhibited varying movement recurrence regions, especially during standard performances, and two exhibited no significant movement recurrence. As the four players who did not follow the dominant movement recurrence pattern are students with less expertise, they were discarded on further analysis.

5. REGIONS OF INTEREST

The 6 expert musicians with similar recurrence patterns were selected for a detailed movement analysis over their

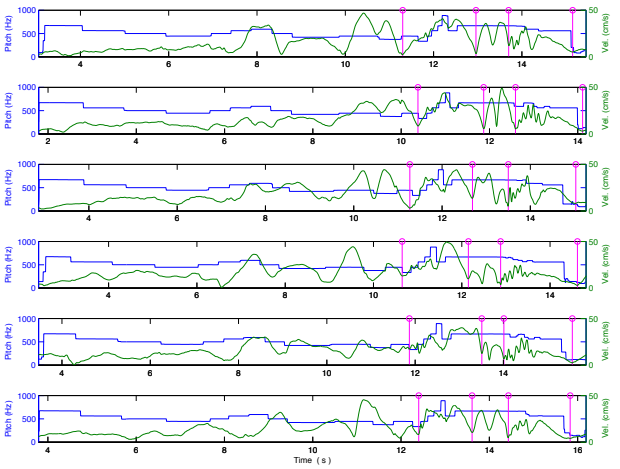

Figure 6. Regions of interest for the 6 performances by Musician 4, and their constituent gestures, marked by the vertical lines. Standard performances on top 3 plots.

high recurrence regions. The movement segments obtained by the tangential velocity minima criteria, and the motion recurrence map analysis were used together to define regions of interest in each of these 6 musicians' performances. These regions of interest were defined along the final part of the excerpt and consist of the movement segments contained inside the high recurrence areas, for each of the performances selected. The 3D spatial trajectory of the clarinet bell was analyzed along each region of interest, in order to group its constituent movement segments into representative physical gestures, based on geometrical and temporal characteristics. Figure 6 illustrates the definition of the regions of interest and their constituent gestures for Musician 4. Figure 7 shows a 2D frontal view of each of their spatial trajectories, according to the reference plane shown in Figure 2.

According to figures 6 and 7, the musician executes recurrent and well defined gestures with the clarinet, inside the region of interest. The number, sequence, geometry and duration of these gestures are all highly recurrent along the performances. Most noticeably, the first gesture of the

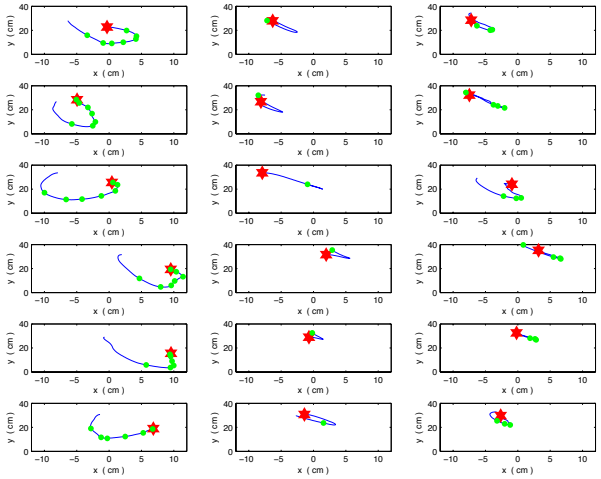

Figure 7. Spatial trajectories of the 3 gestures defined in each region of interest for the performances by Musician 4. The red star marks the initial point and the green circles indicate the note onsets. Each row represents one of the 6 performances.

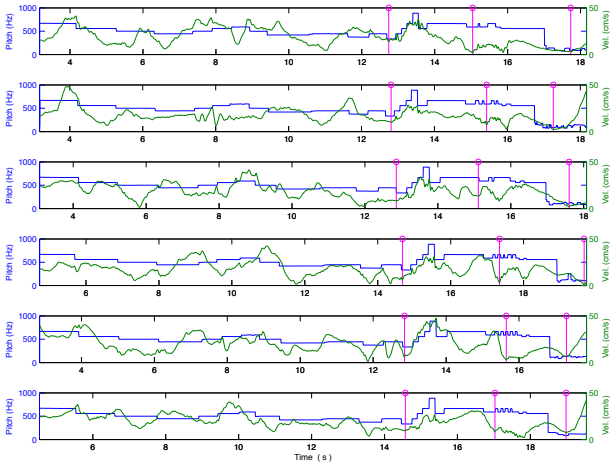

Figure 8. Regions of interest for the 6 performances by Musician 7, and their constituent gestures.

sequence, shown in the left column, always starts at the beginning of the same musical phrase, and consists of a clockwise partial elliptical movement, with significant extension and duration.

This analysis was conducted over the other 5 selected players, with similar results. Figures 6 and 7 show another example. In each case, the musician executed a recurrent sequence of gestures with the clarinet, showing that the tangential velocity based movement segmentation and recurrence analysis, reveals highly representative music related gestures. The occurrence of such recurrent gestures and their strong relation to the musical structure, specially for the most skilled players, constitute strong evidences of the musical meaning of the musician's physical movements and their relevance to music performance.

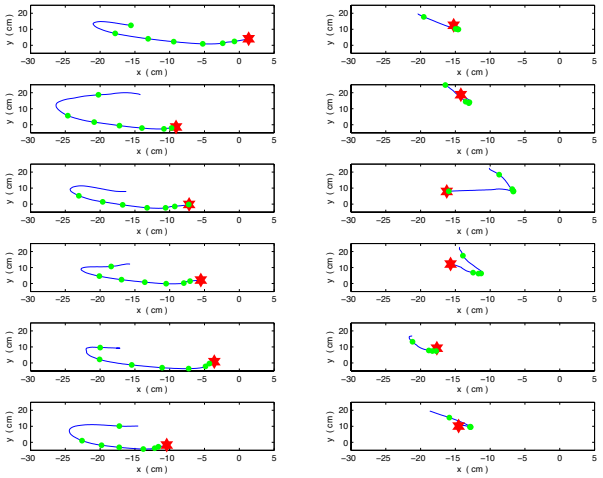

Figure 9. Spatial trajectories of the 2 gestures defined in each region of interest in the performances by Musician 7.

6. RECURRENT GESTURES

The definition of proper regions of interest in the musical performances, and their subdivision into representative and recurrent physical gestures, made possible a further local parametrical analysis on these gestures.

The resulting gestures were subjected to Principal Component Analysis (PCA), in order to investigate the spatial dimensionality of their trajectories. Calculating the percentage of total variance accounted for by the first principal component alone, by the first two principal components, and then by the three components, it is possible to define if the gesture trajectory is mainly unidimentional, mainly bidimentional or tridimensional. The percentage of variance accounted for by the first two principal components in each gesture represents a planarity index for its trajectory, while the percentage of variance accounted for by the first principal component represents a unidimentionality index for its trajectory. The results reveal that all recurrent clarinet gestures are highly planar, with planarity indexes always above 97%, and above 99% in 80% of the cases. They also show that a significant part of these gestures is also highly unidimentional, since 50% of them exhibit unidimentionality indexes above 95%, and 80% exhibit unidimentionality indexes above 80%.

A comparison was also established between the two performance styles, based on three gestural features: the total spatial distance covered along its trajectory (cm), the time duration of the gesture (s), and the mean tangential velocity along its trajectory (cm/s). The results indicate that during standard performances, within the defined regions of interest, the clarinet players execute recurrent clarinet gestures with greater spatial amplitude (28% larger on average) and at larger mean velocities (26% larger on average) than in the metronome controlled performances. The use of the metronome as a control device makes the musicians' gestures and their respective mean velocities smaller in general, but exerts little effect over their time durations, which became only slightly smaller.

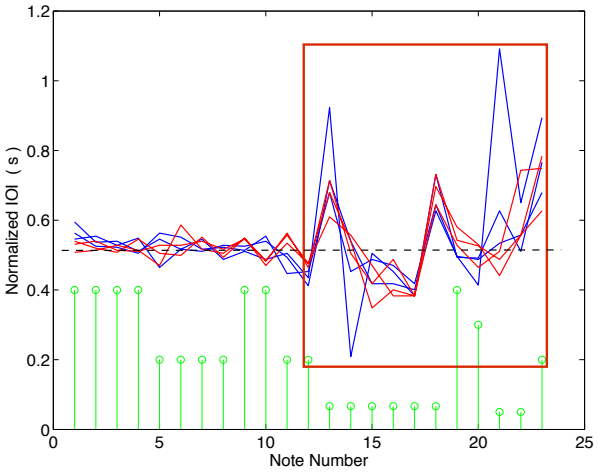

Figure 10. Note intra-onset intervals for Musician 1, normalized relative to a quarter-note. Standard performances in blue, metronome performances in red, relative nominal score value of the notes in green. The black dashed line shows the expected values. The region of interest is highlighted in red.

7. ACOUSTICAL ANALYSIS

The movement analysis indicated that the most significant recurrent clarinet gestures occurred during the final part of the musical excerpt, according to Figure 1. In order to further investigate the importance of this final part of the music, the relative note durations were also analyzed along the performances. To do that, the intra-onset intervals (IOI's) were calculated, based on the note onsets extracted from the audio signals. In order to get a relative idea of the note durations, according to their expected nominal durations in the score, the IOI's were normalized relative to a quarter-note. This way, if all notes were executed precisely with the note durations defined in the score, the plot representing the evolution of note IOI's along the performance would show a straight horizontal line. Any deviation from this expected horizontal line indicates a note duration manipulation by the musician. Figure 10 exemplifies this analysis for Musician 1.

The evolution of the note intra-onset intervals also reveals a great contrast between the initial and final part of the excerpt. In the initial part, there is little manipulation of the note durations by the musician, while in the final part, inside the red rectangle, the musician executes significant and recurrent manipulation over the note durations, indicated by the large deviation of the IOI curves around the black horizontal line. This sort of behaviour was observed for almost all of the musicians in the study. Figure 11 shows another example. This validates the previous assumption of a strong expressive content being imposed by the musicians in this final part of the music, and being reflected in their corresponding clarinet gestures, since both movement and note duration analysis led to related results.

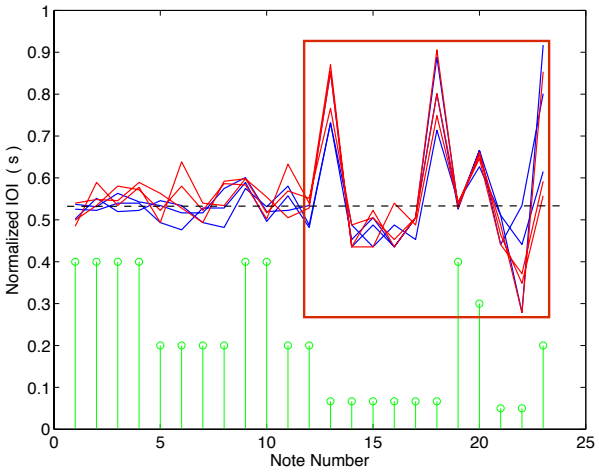

Figure 11. Note intra-onset intervals for Musician 7, normalized relative to a quarter-note.

8. CONCLUSION

The method defined for movement representation, segmentation and recurrence analysis led to representative clarinet gestures, and to preliminary relations between gestural and musical aspects during the performances, and also between the two proposed performance styles, regarding gestural features. A recurrent sequence of planar gestures, more restrained in the metronome cases, was found in each region of interest, which was shown to be a key moment in the music. The strong recurrence and relation to the musical structure found in these physical gestures reinforce the assumption of a musical meaning behind the musicians ancillary movements, related to their expressive intentions and important for the desired musical outcome.

This method will now be applied to a quite larger data set, including new musical excerpts and groups of musicians, aiming at a detailed musical analysis and characterization of individual musicians, musical passages and performance styles.

Acknowledgments

The authors would like to thank the brazilian agencies CNPq, CAPES and FAPEMIG for funding this study.

9. REFERENCES

[1] T. Campolina, M. Loureiro, and D. Mota, "Expan: a tool for musical expressiveness analysis," *Proceedings of the 2nd International Conference of Students of Systematic Musicology*, pp. 24–27, 2009.

[2] M. Loureiro, H. Yehia, H. de Paula, T. Campolina, and D. Mota, "Content analysis of note transitions in music performance," *6th Sound and Music Computing Conference*, pp. 355–359, 2009.

[3] A. Gabrielsson, "Expressive intention and performance," in *Music, Mind and Machine*, Steiner, Ed. New York: Springer, 1995, pp. 35–47.

[4] P. Juslin, "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.

[5] D. Fenza, L. Mion, S. Canazza, and A. Rodà, "Physical movement and musical gestures: A multilevel mapping stategy," *Proceedings of Sound and Music Computing*, 2005.

[6] A. Camurri, B. Mazzarino, M. Ricchetti, R. Timmers, and G. Volpe, "Multimodal analysis of expressive gesture in music and dance performances," *Springer-Verlag Berlin Heidelberg*, 2004.

[7] A. Barbosa, H. Yehia, and E. Vatikiotis-Bateson, "Temporal characterization of auditory-visual coupling in speech," *Proceedings of Meetings on Acoustics*, vol. 1, pp. 1–14, 2008.

[8] L. Naveda and M. Leman, "Representation of samba dance gestures, using a multi-modal analysis approach," *5th International Conference on Enactive Interfaces*, pp. 68–74, 2008.

[9] M. Wanderley, B. Vines, N. Middleton, C. McKay, and W. Hatch, "The musical significance of clarinetists' ancillary gestures: An exploration of the field," *Journal of New Music Research*, vol. 34, no. 1, pp. 97–113, 2005.

[10] B. Vines, C. Krumhansl, M. Wanderley, and D. Levitin, "Cross-modal interactions in the perception of musical performance," in *Cognition*, E. B.V., Ed., 2006, vol. 101, pp. 80–113.

[11] E. Teixeira, M. Loureiro, and H. Yehia, "Methodological aspects of the research in musical expressiveness based on corporal movement information," 2010, unpublished Report.

[12] B. Caramiaux, M. Wanderley, and F. Bevilacqua, "Segmenting and parsing instrumentalists' gestures," *Journal of New Music Research*, vol. 41, no. 1, pp. 13–29, 2012.

[13] F. Desmet, L. Nijs, M. Demey, M. Lesaffre, J. Martens, and M. Leman, "Assessing a clarinet player's performer gestures in relation to locally intended musical targets," *Journal of New Music Research*, vol. 41, no. 1, pp. 31–48, 2012.

[14] N. Rasamimanana, "Towards a conceptual framework for exploring and modelling expressive musical gestures," *Journal of New Music Research*, vol. 41, no. 1, pp. 3–12, 2012.

[15] J. Davidson, "Visual perception and performance manner in the movements of solo musicians." *Psychology of Music*, vol. 21, pp. 103–113, 1993.

[16] J. W. Davidson, "What type of information is conveyed in the body movements of solo musician performers?" *Journal of Human Movement Studies*, vol. 6, pp. 279–301, 1994.

[17] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movements," *Music Perception*, vol. 24, no. 5, pp. 433–454, 2007.

[18] A. Barbosa, R. Déchaine, E. Vatikiotis-Bateson, and H. Yehia, "Quantifying time-varying coordination of multimodal speech signals using correlation map analysis," *Journal of the Acoustical Society of America*, vol. 131, no. 3, pp. 2162–2172, 2012.

[19] P. Senin, "Dynamic time warping algorithm review," Information and Computer Science Department, University of Hawaii at Manoa, Honolulu, USA, Tech. Rep., 2008.