TESE DE DOUTORADO Nº 186

MOTION ANALYSIS OF CLARINET PERFORMERS

Euler da Cunha Francisco Teixeira

DATA DA DEFESA: 03/07/2014

Universidade Federal de Minas Gerais

Escola de Engenharia

Programa de Pós-Graduação em Engenharia Elétrica

MOTION ANALYSIS OF CLARINET PERFORMERS

Euler da Cunha Francisco Teixeira

Tese de Doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do Título de Doutor em Engenharia Elétrica.

Orientador: Prof. Hani Camille Yehia
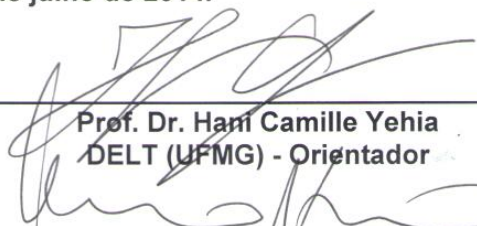
Belo Horizonte - MG

Julho de 2014

"Motion Analysis of Clarinet Performers"

Euler da Cunha Francisco Teixeira

Tese de Doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Doutor em Engenharia Elétrica.

Aprovada em 03 de julho de 2014.

Por:

_____
Prof. Dr. Hani Camille Yehia
DELT (UFMG) - Orientador

_____
Prof. Dr. Maurício Alves Loureiro
Escola de Música (UFMG) - Coorientador

_____
Prof. Dr. Adriano Vilela Barbosa
DELT (UFMG)

_____
Prof. Dr. Sérgio Freire Garcia
Escola de Música (UFMG)

_____
Prof. Dr. Jônatas Manzolli
Departamento de Música (UNICAMP)

_____
Prof. Dr. Marcelo Mortensen Wanderley
CIRMMT (McGill University)

Motion Analysis of Clarinet Performers

Euler C. F. Teixeira

Doctoral Thesis - PPGEE/UFMG

July 25, 2014

Abstract

Musical expressiveness is a concept that is difficult to formalise by quantitative data and its analysis usually relies on some sort of subjective evaluation. Today there is a growing interest in methods and cues used to extract, quantify, analyse and synthesise these expressive intentions. This have been done mainly through the audio analysis of music performances, identifying the acoustic parameters capable of describing their expressive content.

This study expands acoustic analysis methods for investigating the expressive intentions of musicians, incorporating information about their body movements during musical performances. It presents a method to define and analyse the physical gestures executed by the musicians while playing their instruments, and to extract motion parameters that can be quantitatively related to their expressive intentions and to the musical structure.

The gesture consistency of 13 clarinetists is evaluated during several performances, establishing an objective relation between their expressive gestural patterns and the music structure of two selected excerpts, by Mozart and Brahms. A method is defined to represent, segment and analyse the patterns of recurrence on motion data during musical performances. Recurrent physical gestures were extracted during clarinet performances and analysed based on gestural features, comparing different musicians, musical passages and experimental conditions. Results indicate recurrent sequences of clarinet gestures in regions of the excerpts that were shown to be related to key musical moments. A corresponding analysis is conducted over the acoustic data, searching for related parametrical patterns that could validate the results of the motion analysis.

The information obtained can be used to define an integrated method to parametrise and quantify the expressive intentions of musicians. This method could be incorporated to musical synthesis, recognition, analysis and teaching systems, or used in theoretical studies in musicology, human cognition and physiology, ultimately defining a musical meaning for the physical gestures of musicians during their performances.

Resumo

A expressividade musical é um conceito difícil de ser formalizado com base em dados objetivos e sua análise geralmente depende de algum tipo de avaliação subjetiva. Atualmente há uma crescente busca por métodos e descritores capazes de extrair, quantificar, analisar e sintetizar essas intenções expressivas. Isso vem sendo feito principalmente através da análise de áudio das execuções musicais, identificando os parâmetros acústicos capazes de descrever o seu conteúdo expressivo.

Esse estudo expande os métodos de análise acústica para a investigação das intenções expressivas dos músicos, incorporando informações sobre seus movimentos corporais durante as execuções musicais. Uma metodologia é proposta para a definição e a análise dos gestos físicos realizados por eles ao tocarem, e para a extração de parâmetros de movimento que possam ser relacionados objetivamente às suas intenções expressivas e à estrutura musical.

A consistência gestual de 13 clarinetistas é avaliada durante diversas execuções musicais, estabelecendo uma relação objetiva entre seus padrões gestuais expressivos e a estrutura musical de dois excertos de Mozart e Brahms. Um método é definido para representar, segmentar e analisar os padrões de recorrência nos dados de movimento das execuções. Gestos físicos recorrentes foram extraídos durante execuções de clarineta e analisados com base em parâmetros gestuais, comparando diferentes músicos, passagens musicais e condições experimentais. Os resultados indicam sequências recorrentes de gestos de clarineta em regiões dos excertos relacionadas a momentos de alta relevância musical. Uma análise correspondente é conduzida sobre os dados acústicos, buscando padrões paramétricos relacionados, capazes de validar os resultados da análise de movimento.

As informações obtidas podem ser usadas para definir um método integrado de parametrização e quantificação das intenções expressivas dos músicos. Esse método pode ser incorporado a sistemas de síntese, reconhecimento, ensino e análise de música, ou utilizado em estudos teóricos nas áreas de musicologia, cognição e fisiologia humanas, possibilitando a definição de um significado musical para os gestos físicos de músicos durante execuções musicais.

Acknowledgements

Contents

List of Figures

4

6

Chapter 1

Introduction

With the development of new interfaces for music composition, performance and learning, there is a growing interest in methods that could be used to describe and analyse musical expressiveness according to a set of quantitative parameters (Gabrielsson, 2003). The expressive content is responsible for the human and individual character observed in the performances of each musician and it is fundamental for the desired listener's and performer's experience. Musical synthesis systems are already capable of executing music pieces based on a musical score or on MIDI files with great accuracy, but it is hard to effectively incorporate this expressive musical content to such systems and so the resulting synthesised pieces still sound unnatural to most listeners.

This subject has been studied by a wide group of researchers, such as musicologists, psychologists, neuroscientists and engineers. Even so, the concept of musical expressiveness is still hard to formalise by quantitative data. Research in this direction has been done mainly through the audio analysis of musical performances, by extracting musical content information, such as energy envelope, pitch, note onsets and offsets, attack duration and spectral centroid, directly from the acoustic data. This methodology was used by Campolina, Loureiro, and Mota (2009), De Poli,

Canazza, Drioli, Rodà, and Vidolin (2004), and Maestre and Gómez (2005), among others, to conduct quantitative analysis of musical performances. Studies of this sort have shown that musicians make use of small deviations, regarding note durations, articulations, intensity, pitch and timbre, in order to convey their musical intentions, according to previous works by Gabrielsson (1995) and Juslin (1997, 2000).

This study expands acoustic analysis methods for investigating the expressive intentions of musicians, incorporating information about their body movements during musical performances. A method is presented to define and analyse the physical gestures executed by the musicians while playing their instruments, and to extract motion parameters that can be quantitatively related to their expressive intentions and to the musical structure, as suggested by Camurri, De Poli, Leman, and Volpe (2001); Camurri, Mazzarino, Ricchetti, Timmers, and Volpe (2004); Fenza, Mion, Canazza, and Rodà (2005) and Leman (2007). Multi-modal investigation has also been successfully employed in studies related to the analysis of speech, by Barbosa, Yehia, and Vatikiotis-Bateson (2008) and to the analysis of dance, by Naveda and Leman (2009), examining the coupling between their acoustic 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 expressive intentions of the musician in a particular performance, as shown by Wanderley, Vines, Middleton, McKay, and Hatch (2005) and Vines, Krumhansl, Wanderley, and Levitin (2006b). The movements that are part of a performance, but that are not produced for the purpose of sound generation, have been usually designated as ancillary gestures (Wanderley et al., 2005) or expressive movements (Davidson, 1993).

Recent studies have searched for methods to extract and analyse these movements in detail. In (Teixeira, Loureiro, and Yehia, 2010) a tangential velocity based model was proposed to represent, segment and analyse expressive movements based on local gestural parameters. Caramiaux, Wanderley, and Bevilacqua (2012) used a Hidden Markov Model (HMM) to segment movements into primitive shapes, selected from a base dictionary. Desmet, Nijs, Demey, Lesaffre, Martens, and Leman (2012) proposed another statistical movement segmentation procedure and pointed to subjective links between these segments and the musical score. Rasamimanana (2012) defined a conceptual framework considering performer-instrument relationships that can provide ground to model expressive gestures using a space of possibilities.

These studies have not yet revealed a clear objective relation between gestural and musical patterns in the performances, and there is still no quantitative method that can be widely used to extract and analyse such information from the motion capture data. However, there is strong evidence that such expressive information is present in the body movements of musicians, providing valuable information to better comprehend expressiveness from a multimodal point of view. Davidson (1993, 1994, 1995) was able to observe in a subjective evaluation that the image of a piano performance can be as important as the sound in perceiving the musician's expressiveness, by presenting to subjects its visual and acoustic components separately.

According to Camurri et al. (2004) and Dahl and Friberg (2007), three steps are necessary in order to identify relations between the body movements of performers and their expressive musical intentions. The first step is to track points of interest in the body and instrument of the musician during several musical performances, searching for patterns of temporal and spatial evolution, in order to define significant and recurrent physical gestures. After that, the gestures of

10

different performers and musical passages must be compared, taking into account spatial, temporal and musical parameters. The final step is to conduct an analysis over the corresponding musical and acoustic data, searching for related parametric patterns coupled with the motion analysis. With this method, it is possible to investigate where the expressive content can be found in the body movements of musicians and how it relates to the music, ultimately defining a musical significance for the physical gestures of musicians during performances.

This study presents a procedure to extract physical gestures of clarinet players during performances of pieces from the classical repertoire, based on movement segmentation and analysis of recurrence patterns, analysing their musical location and gestural features. Movement tracking and segmentation have been addressed in recent studies, but recurrence is still a less investigated topic. The gesture consistency of several clarinetists will be evaluated, relating their expressive gestural patterns to the music structure of excerpts of masterpieces in the clarinet repertoire and also to expressive acoustic patterns in the performances. The main hypothesis is that the ancillary physical gestures employed by expert musicians during performances are closely related to their musical intentions towards the musical realisation, and that the strong expressive intentions imposed in key regions of the music structure are thus reflected in these gestures. The results of a movement and an audio analysis of several clarinet performances will be analysed and correlated in order to quantitatively evaluate this hypothesis.

All the definitions regarding the experiments used in the study, such as the selected group of performers, corporal points of interest, musical pieces, experimental conditions and data collection procedures, as well as the computational tool developed for the movement and audio analysis, will be described and discussed in detail further in the text, along with the obtained results. The text is structured in five chapters, besides this introduction, describing respectively, the

experiments conducted for the study, the movement analysis methodology used, the analysed results, a discussion regarding the results and finally a general conclusion, including future perspectives of the research.

Chapter 2

Experimentation

2.1 Experimental Methodology

The objective of the study is to analyse the expressive content of musical performances by a group of clarinet players based on the corresponding motion and acoustic data. The analysis is based on two experiments.

2.1.1 First Experiment

In the first experiment ten classical clarinet players performed a short excerpt of six bars extracted from the first movement of the Quintet for Clarinet and Strings in A Major, Kv 581 by W. A. Mozart [1] (Figure 2.1). The musicians were asked to play according to two distinct experimental conditions. First, normally expressive performances, as in a real concert situation (standard condition). After that, the player was asked to follow a metronome, set to a tempo estimated from

---

[1]Mozart wrote his Clarinet Quintet in 1789, for the clarinetist Anton Stadler. Stadler was reputed to be able to produce a voice-like tone, with exceptional mastery of the low register of his basset clarinet in A, for which the quintet was originally written. One of the earliest and best-known works written especially for the instrument, it is usually played on standard soprano clarinet in A. It remains exceptionally popular today due to its lyrical melodies and it is a highly representative work of the Classical period of western music in the clarinet repertoire.

his/her previous standard performances (metronome condition). Each of the ten clarinet players performed this excerpt six times without accompaniment, three performances for each of the two experimental conditions.

Figure 2.1: Main theme from first movement of Mozart's Quintet for Clarinet and Strings in A Major, Kv 581, performed in the first experiment.

The basic assumption behind the definition of these experimental conditions is that while following the metronome the attention of the musicians would be directed primarily towards an external control device, what should inhibit or alter their natural expressive intentions during the musical performances. On the other hand, during standard performances, without the metronome reference, the musicians are able to perform the musical piece freely, manipulating the musical notes and also their body movements in order to achieve the desired musical outcome. Based on these assumptions and these two experimental conditions, the goal is to obtain an objective way to compare between free musical performances and controlled performances restrained by the metronome beat. This is a good way to unveil specific aspects related to the expressive intentions of musicians, based on referential performances, as previously reported in (Teixeira et al., 2010).

2.1.2   Second Experiment

In the second experiment 13 classical clarinet players (including the ten participants of the first experiment) performed a longer phrase, with much more complex harmonic progressions and melodic lines, that has also been used in other similar studies by Wanderley (2002), among others. The excerpt was extracted from the first movement of the Sonata for Clarinet and Piano in F minor, op. 120 No. 1 by

Johannes Brahms [2] (Figure 2.2). Each of these clarinetists performed this excerpt three times without accompaniment, in the standard experimental condition only.

Figure 2.2: Excerpt from Brahms' Clarinet Sonata No.1 in F minor, Op.120 No.1, performed in the second experiment.

2.2   Data Acquisition

In order to perform the data analysis, an adequate motion capture strategy must be used to track the body movements of musicians during the recording of several musical performances. Motion tracking was done with high-end 3D motion capture devices, the NDI Optotrak Certus and the NDI Optotrak 3020. These devices consist of a tracker built with three infra-red cameras positioned along one axis, tracking the spatial position of active infra-red LED markers inside a tridimensional measurement volume, together with synchronous audio recording, which is

---

[2]Brahms wrote his two Clarinet Sonatas, op. 120 Nos. 1 and 2, in 1894, for the clarinetist Richard Mühlfeld. The form of the clarinet sonata was still largely undeveloped until the completion of these sonatas, after which the combination of clarinet and piano was more readily used by later composers. These were the last chamber pieces Brahms wrote before his death. Together with his Clarinet Trio and Clarinet Quintet, also written for Mühlfeld shortly before, these sonatas are considered to be two of the greatest masterpieces in the clarinet repertoire, being highly representative of the Romantic period of western music.

fundamental for the objectives of the study. High sampling rates are possible with a large number or markers and a spatial resolution of 0.1 millimetre.

During data acquisition the clarinetists performed the excerpts standing up, sideways to the Optotrak tracker. Motion capture markers were placed on their bodies and instruments. Multiple mechanically coupled markers were used on the head and clarinet of the musician, allowing for their definition as rigid bodies. The motion capture tracker was placed vertically, two meters away from the subject, in order to provide an adequate measurement volume. Figure 2.3 shows the data acquisition setting, as seen by the Optotrak tracker (xy plane). Motion was captured at a sampling rate of 100 frames per second. Audio was recorded synchronously at a sampling rate of 44.1kHz using a condenser microphone positioned one meter away from the clarinet. The performances were also filmed with a consumer video camera for future reference.

Figure 2.3: Data acquisition setting, from the motion capture tracker viewpoint (xy plane). An arrow indicates the positioning of the clarinet bell marker.

Chapter 3

Movement Analysis Methodology

In this chapter the movement analysis methodology developed in this study is described in detail, using the performances of Mozart's excerpt by one of the clarinetists as an example. Figure 3.1 presents a schematic diagram of the procedure.
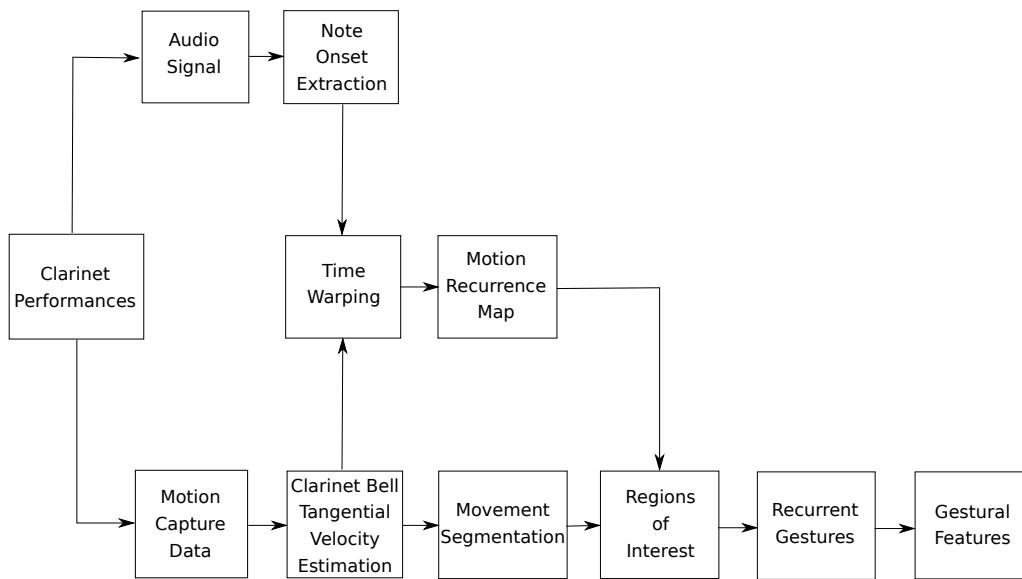
Figure 3.1: Schematic diagram illustrating the movement analysis methodology. Note onsets are extracted from the audio signal and the clarinet bell tangential velocity is estimated from the motion capture data. These are used to compute a motion recurrence map, with the aid of time warping, in order to define regions of interest. Clarinet bell movements are segmented and then recurrent gestures are defined in these regions, from which gestural features are extracted.

3.1 Data Preprocessing

Data recorded during the 99 performances, 60 in the first experiment (10 players) and 39 in the second experiment (13 players), was inspected for occlusion and exported from the Optotrak data acquisition platform into MatLab matrices and vectors that are used in the analysis stage. The audio vector of each performance was processed using the EXPAN system (Campolina et al., 2009), developed in UFMG School of Music, to extract its pitch and energy envelope curves, from which all note onsets and offsets used in the analysis were detected.

3.2 Movement Representation

Movement analysis in this study will be based on the clarinet bell motion. The clarinet motion has been the object of previous studies (Wanderley, 2002; Wanderley et al., 2005; Caramiaux et al., 2012) and it is believed to be an important indicator of expressive movements made by the musician.

In this study the movement of the clarinet bell is taken relative to a static reference (instead of a dynamic one, such as the mouthpiece), the origin of the cartesian coordinate system, located in the centre of the Optotrak tracker. This way the clarinet bell movement incorporates any motion performed by the musicians with their feet, knees, torso, neck and arms, and can thus be seen as a general indicator of the players' movements. Optical flow techniques were already used to define a general motion indicator (Barbosa et al., 2008), but the method presented in this study also allows the precise analysis of a specific point (a single marker placed at the clarinet bell), well known for its expressive character, including its 3D trajectory, in order to define recurrent gestures and many associated gestural features.

In order to analyse the evolution of the tridimensional motion of the clarinet bell in conjunction with the acoustic data, using signal processing techniques, an effective scalar representation of the motion data in time is needed. A simple and adequate solution in this case is to use the tangential velocity of the clarinet bell marker's trajectory, estimated using the Euclidian distance between the positions of this marker in two subsequent samples and the sampling rate f_s, according to the following expression:

$$v_t(n) = f_s\sqrt{\big(x(n+1) - x(n)\big)^2 + \big(y(n+1) - y(n)\big)^2 + \big(z(n+1) - z(n)\big)^2}.$$

$$(3.1)$$

Figure 3.2: Movement representation for a performance of Mozart's excerpt. The thick grey curve shows the pitch (Hz) and the black curve shows the tangential velocity of the clarinet bell (mm/s).

This unidimensional parameter captures a large amount of information from the musician's movements, as previously demonstrated in a related study (Teixeira et al., 2010). Figure 3.2 illustrates this representation for a performance of Mozart's excerpt. The thick grey curve represents the pitch extracted from the audio signal, with each step representing a musical note in the clarinet melody. The black curve represents the tangential velocity of the clarinet bell extracted from the 3D

19

motion capture data, after low-pass filtering for smoothing. In the figure time is shown in bars, with the numbers indicating the starting point of each bar. This representation provides a general idea of the evolution of the musical performance and the movements employed by the musician during the execution.

3.3   Movement Segmentation

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 acoustic data into musical notes and phrases. This is very useful to analyse many aspects of the musical performance according to local parameters. It is also very important to develop a procedure to segment the movement data accordingly. Unlike the acoustic data, there are no basic units established to segment the movements into, but it is possible to divide those movements into representative segments, based on their geometrical and temporal attributes.

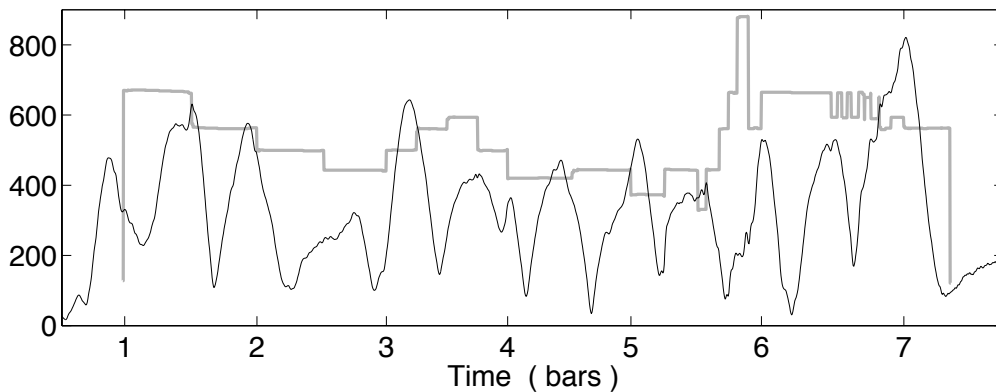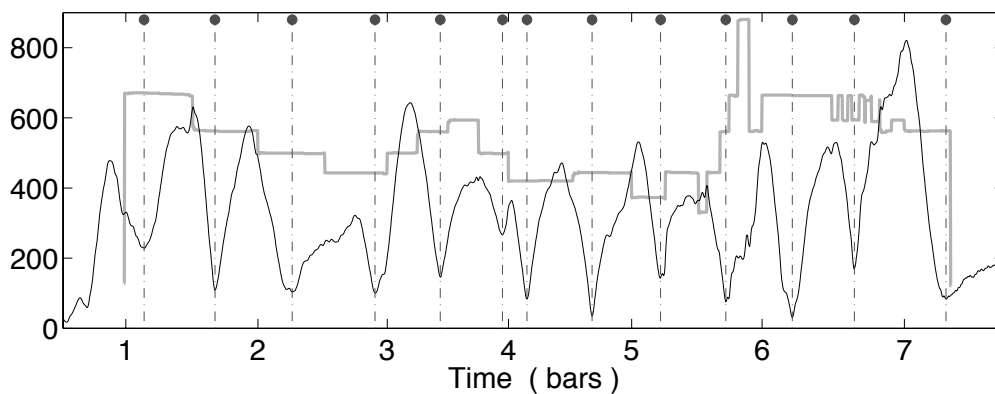Figure 3.3: Movement segmentation for a performance of Mozart's excerpt. The thick grey curve shows the pitch (Hz), the black curve shows the tangential velocity of the clarinet bell (mm/s) and the vertical dashed lines mark the movement segments.

This can be done based on the tangential velocity representation of movement illustrated in Figure 3.2, assuming that its local minima correspond to inflection

points of the musician's movement, where the motion direction or character is most likely to suffer a sudden change (Teixeira et al., 2010). These points were thus used as a basis for the segmentation procedure, defining movement segments between subsequent local minima in the tangential velocity curve. Figure 3.3 illustrates this segmentation method for the performance of Mozart's excerpt shown in Figure 3.2. The procedure is applied to all of the performances analysed in this study.

3.4   Movement Recurrence

In order to relate the movements of the musicians to their expressive intentions and to the musical structure, the recurrence of movements within the excerpt over different performances by the same musician is analysed. The interpretation of well known musical passages, such as those used in this study, by an expert player, arouses from a long process of learning and practice, during which the musician acquires not only a great familiarity with the music, but also a sedimentation of the way he or she plays it, based on a quite personal interpretation concept. It can thus be expected that many aspects of his/her musical intentions remain mostly unchanged over different performances of an excerpt. If any of the player's ancillary gestural pattern is related to these musical intentions, and also acquired during this process, it is also expected to be recurrent over consecutive performances. Previous studies by Caramiaux et al. (2012) and Wanderley et al. (2005) have already investigated gestural motifs and repetitive movements. This study presents a systematic method to highlight the recurrent gestures in the performances.

In order to search for recurrent patterns in the movements of the musicians an instantaneous correlation algorithm proposed by Barbosa, Déchaine, Vatikiotis-Bateson, and Yehia (2012) will be used over the motion data. This algorithm was recently developed in the University of British Columbia in Vancouver, Canada,

in collaboration with UFMG, and its application for motion recurrence analysis is explained in detail in this section. The complete description and mathematical formulation of the algorithm can be found in the paper by Barbosa et al. (2012). 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 bidimensional correlation map between the two signals, similar to the plot shown in the bottom half of Figure 3.4. The horizontal axis in this figure represents time, from the first note onset to the last note offset of the clarinet melody, and the vertical axis represents the time offset between the two analysed signals, from -0.5 to +0.5 seconds. The white areas on the map correspond to low correlation values between the signals, while the dark areas correspond to high correlation values.

In this study the correlation map will be used to measure the overall recurrence over a group of motion signals, instead of the correlation between two signals. In order to do that, a correlation map will be calculated for each of the possible signal pairs in the group, and then all the resulting maps will be summed. Since the focus is on recurrence, and thus on the positive correlations, each map will have its negative correlation values truncated to zero before the summation.

For each musician, a correlation map was calculated for all possible pairs of clarinet bell tangential velocity signals (15 in this case, for 6 performances)[1]. To ensure a proper temporal alignment between the signals, in accordance with the musical structure of the excerpt, the velocity curves were time-warped (Senin, 2008) using the note onsets as reference points in the timing model. These correlation maps had their negative values truncated to zero and were then summed and normalised to one, generating a resulting map that provides a recurrence measure for that musician's clarinet bell movement over his/her performances. In order

---

[1]The number of possible signal pairs in a group is given by the combination of its n elements taken two at a time, equal to: $n(n-1)/2$

to highlight regions of interest, of high recurrence, an empirical threshold was applied to the recurrence map, removing values below 0.75. Values below 0.75 were replaced by 0 (represented in white) and values between 0.75 and 1 were kept unchanged (represented in grey scale). The temporal sensitivity parameter η, used in the calculation of instantaneous correlations, was set to 0.05.
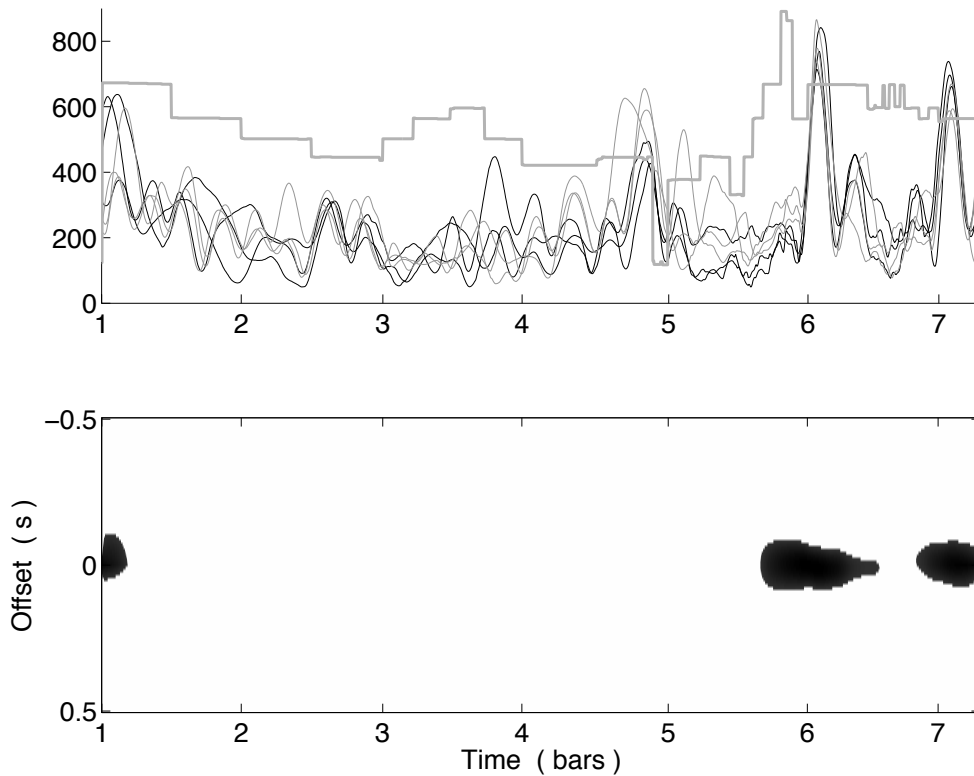
Figure 3.4: Clarinet bell motion recurrence map for the six performances of Mozart's excerpt by Musician 9. Top plot: tangential velocity curves (mm/s), standard performances in black, metronome performances in grey, pitch (Hz) in the thick grey curve. Bottom plot: motion recurrence map, dark regions indicate high recurrence.

Figure 3.4 illustrates the result of this recurrence map analysis for Musician 9. In the top plot of the figure the 6 time-warped clarinet bell tangential velocity curves for this musician are displayed. The 3 standard performances are represented in black and the 3 metronome guided performances are represented in grey. The acoustic pitch is also shown in the thick grey curve for musical reference

23

(mean value for each note, obtained through median filtering). The bottom plot of the figure shows the corresponding motion recurrence map for this musician, relative to his 6 performances of Mozart's excerpt, with dark regions indicating high recurrence.

Analysing Figure 3.4 it is possible to see the regions in this Mozart's theme where this musician employs recurrent movements along his performances. Most noticeably at the last three bars, where all the velocity curves at the top plot are highly correlated. This is confirmed by the large dark areas in the corresponding region of the recurrence map. This motion recurrence map analysis was conducted for the 13 clarinet players in both experiments and the detailed results are presented in the next chapter.

3.5   Regions of Interest

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 the performances by each musician. These regions consist of the movement segments in each performance that are contained inside the high recurrence areas. As an example, Figure 3.5 shows in detail the region of interest seen at the end of Mozart's excerpt in Figure 3.4, for one of the performances by Musician 9. It corresponds to the last three bars of the excerpt, presented in Figure 4.3.

The 3D trajectory of the clarinet bell was then analysed along the region of interest, in order to group its constituent movement segments into representative physical gestures, based on visual inspection of their geometrical characteristics. As already shown in previous studies (Teixeira et al., 2010) most of these gestures have some sort of elliptical trajectory along one plane, with a return point that corresponds to a velocity minima. This indicates round-trip gesture trajectories

24

that occur around an equilibrium position, especially for the most prominent gestures, according to theories by McNeill (2007). Figures 3.5 and 3.6 illustrate how these movement segments are grouped into gestures.
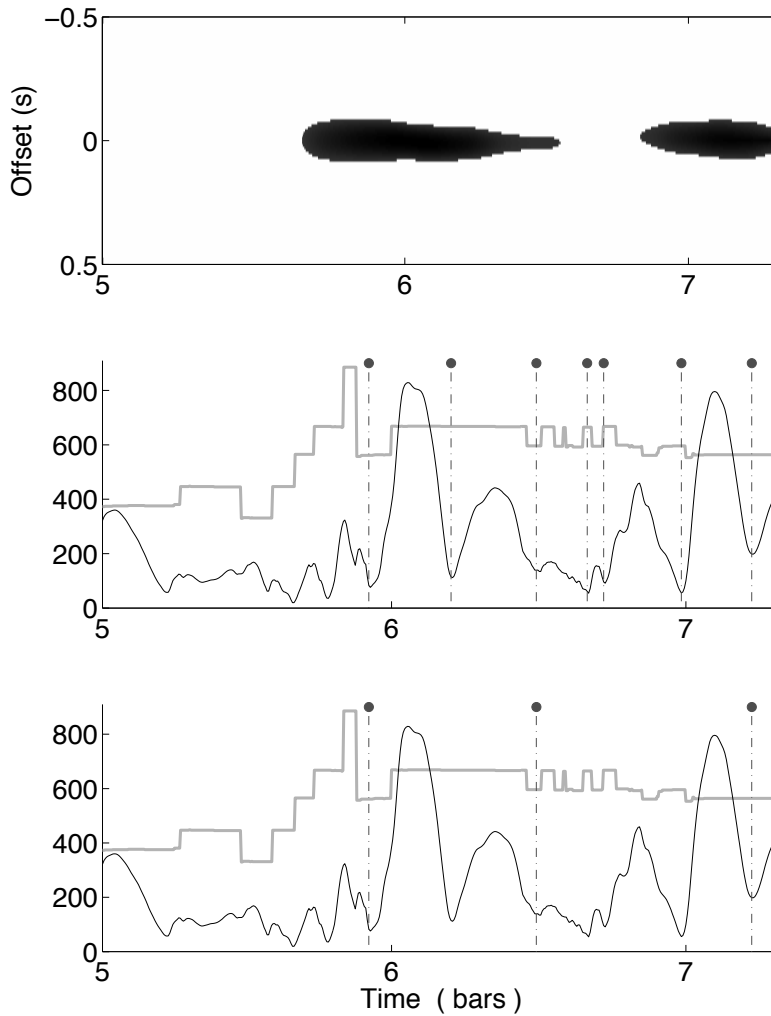
Figure 3.5: Region of interest at the end of Mozart's excerpt for a performance by Musician 9. Top plot: motion recurrence map. Centre plot: location of the movement segments. Bottom plot: location of the resulting gestures. The grey curve shows the pitch (Hz), the black curve shows the tangential velocity (mm/s) and the vertical dashed lines mark the movement segments or the gestures.

The centre plot in Figure 3.5 indicates the location of the six movement segments inside this region of interest. The top plot in Figure 3.6 shows the 3D trajectory of the clarinet bell during the whole region of interest, with the initial

25

points of each movement segment indicated by a star. The inspection of this 3D plot reveals clearly the occurrence of two gestures in this region of interest. The first gesture consists of the first two movement segments and the second gesture consists of the last four movement segments. The gestures defined by this grouping of segments can be seen in the bottom plot of Figure 3.6, and their location in the region of interest is indicated in the bottom plot of Figure 3.5.
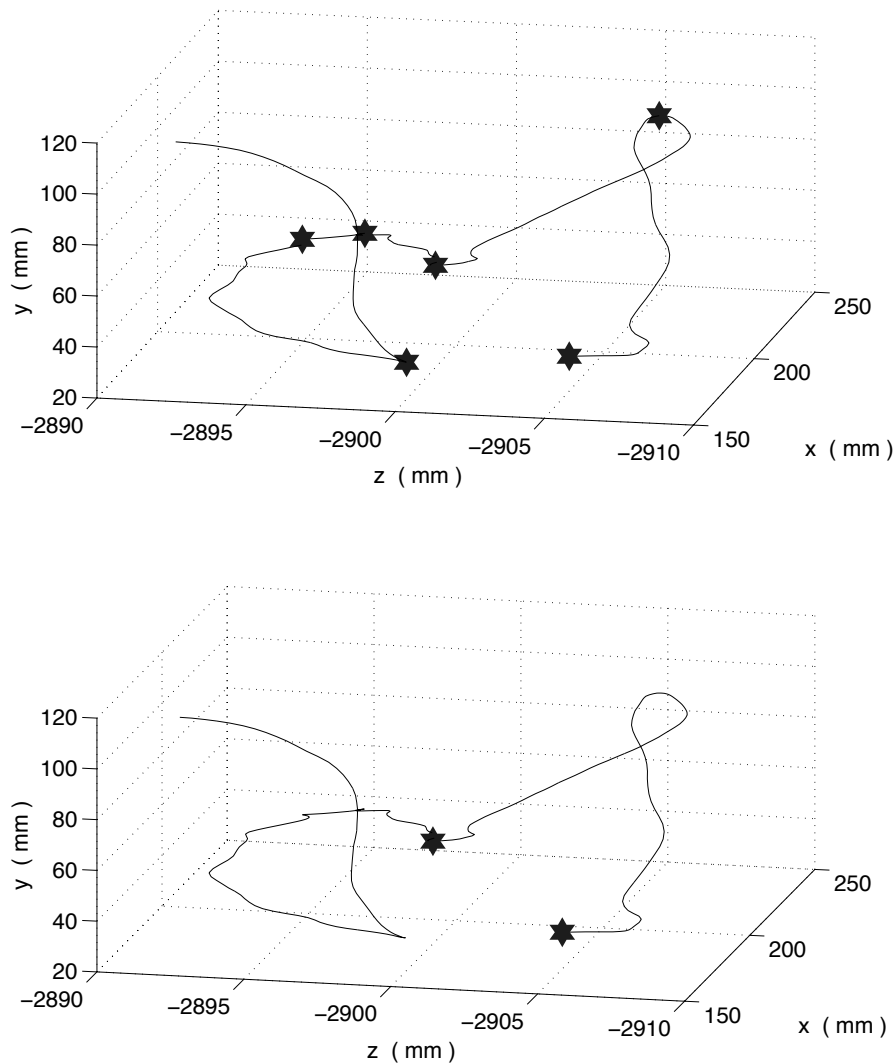
Figure 3.6: 3D trajectory of the clarinet bell in the region of interest at the end of Mozart's excerpt for a performance by Musician 9. The stars mark the initial point of each movement segment in the top plot and the initial point of each resulting gesture (after grouping) in the bottom plot.

This grouping pattern for the movement segments inside the region of interest is then applied to all performances by this musician, in order to define recurrent gestures. As an example, Figure 3.7 shows the location of the recurrent gestures inside this region of interest for the three standard performances by this musician. Figure 3.8 shows a 2D view of each of their trajectories (xy plane, as shown in Figure 2.3), with the grey circles indicating the instants of note onsets.
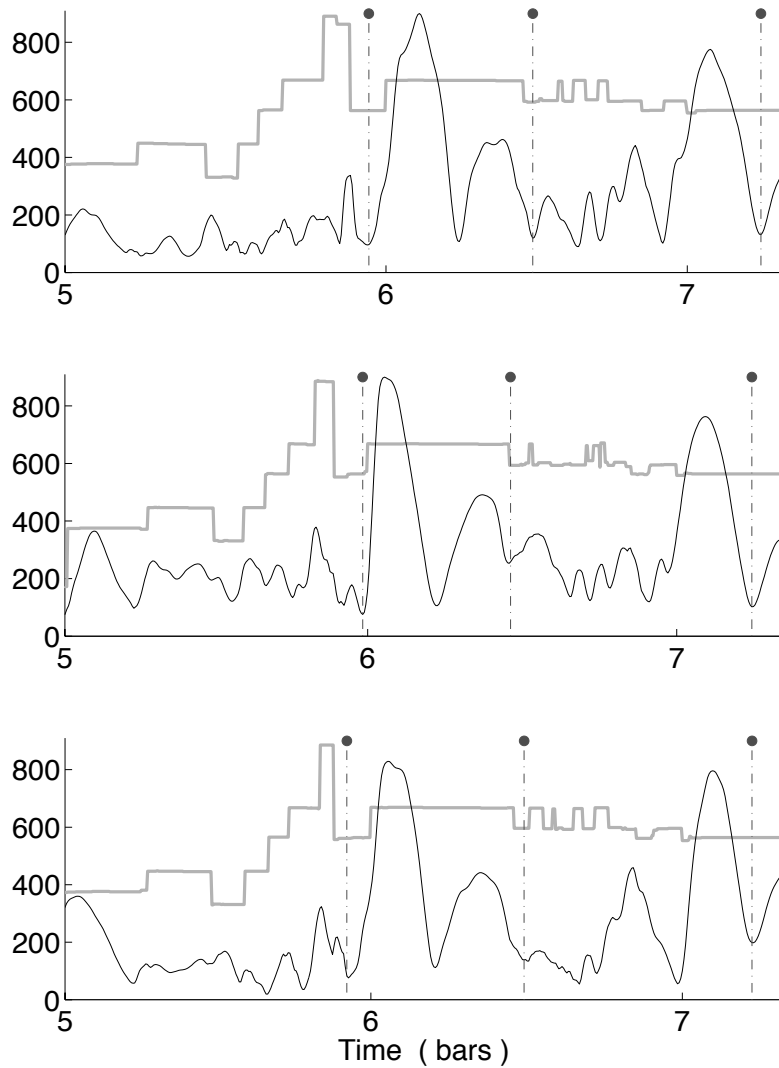
Figure 3.7: Location of the gestures in the region of interest at the end of Mozart's excerpt for the standard performances by Musician 9, marked by the vertical dashed lines. The grey curve shows the pitch (Hz) and the black curve shows the tangential velocity (mm/s).

27
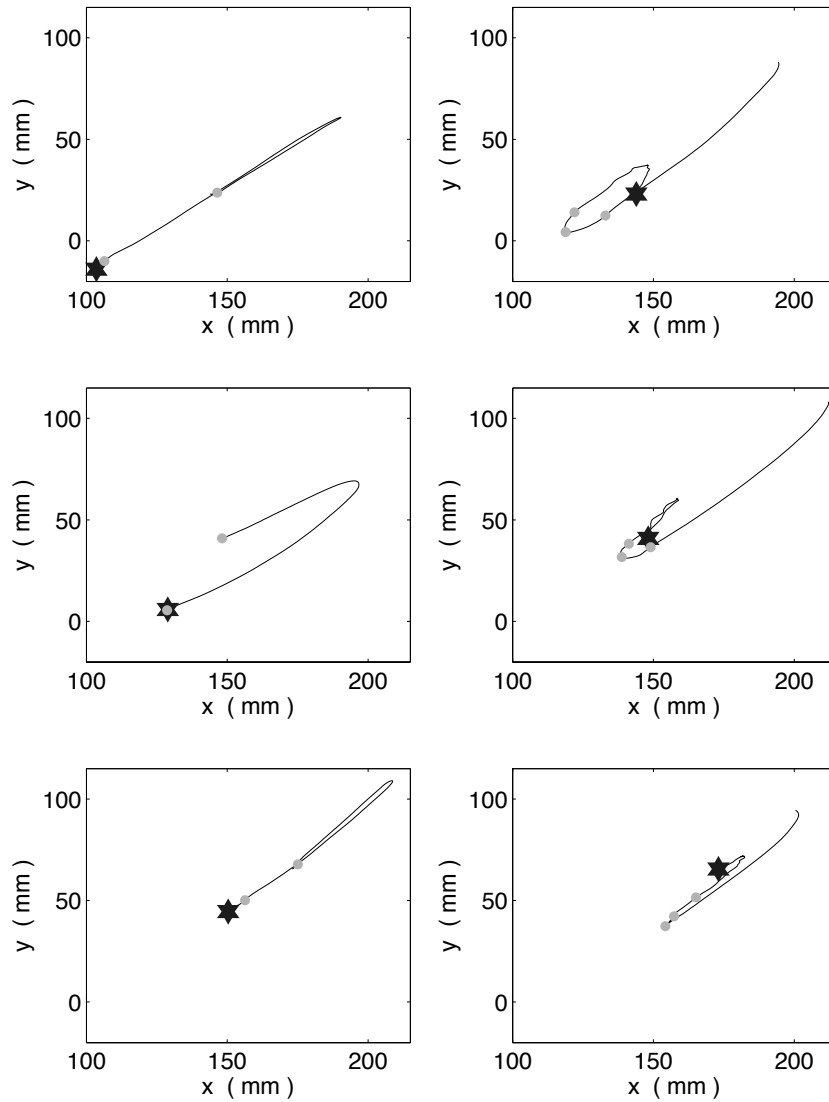
Figure 3.8: 2D trajectories of the two gestures defined in the region of interest at the end of Mozart's excerpt for the three standard performances by Musician 9. The star marks the initial point and the grey circles indicate the note onsets.

As shown in figures 3.7 and 3.8, the musician executes recurrent and well defined gestures with the clarinet bell inside the region of interest. The number, sequence, location, geometry and duration of these gestures are all highly recurrent along the performances. This procedure to define and analyse regions of interest was conducted for all musicians in the study, with results presented in the next chapter.

3.6 Gestural Features

With the definition of regions of interest in the performances, and their subdivision into representative and recurrent physical gestures, it is possible to perform a local parametrical analysis on these gestures. This is done through the extraction of spatial, temporal and statistical features of the defined gestures.

It is important to say that even though time warping was used to compute the motion recurrence maps and locate the regions of interest in the score, the movement segments and resulting gestures were defined on the original unwarped data, in order to preserve the individual expressive aspects of each performance throughout the analysis.

3.6.1 Gestural Dimensionality

The resulting gestures were subjected to Principal Component Analysis (PCA), in order to investigate the spatial dimensionality of their trajectories. The PCA rotates and translates the coordinate system in order to align the first coordinate with the largest data variance, the second coordinate (orthogonal to the first) with the second largest data variance, and so on. This way it is possible to describe the data with the minimum number of dimensions possible, in order to account for the desired amount of the data variance. The PCA also provides the amount of variance accounted for by each principal component. Calculating the percentage of total variance accounted for by the first principal component, by the first two principal components, and then by the three components, makes it possible to define if the gesture trajectory is mainly unidimensional, mainly bidimensional (planar) or tridimensional.

This is the objective of the analysis. The 3D trajectory of each recurrent gesture was subjected to PCA, and the percentage of the total variance accounted for by

the first principal component alone, and by the first two principal components together, were stored in vectors for statistical analysis. The percentage of variance accounted for by the first two principal components in each gesture is used to define a planarity index for its trajectory, while the percentage of variance accounted for by the first principal component is used to define an unidimensionality index for its trajectory.

3.6.2 Gestural Prominence

In order to establish comparisons between particular cases and the two proposed experimental conditions, regarding gestural prominence, each of the recurrent clarinet gestures had three temporal and spatial features extracted from its trajectory. The selected features were: the total spatial distance covered along its trajectory (mm), the time duration of the gesture (s), and the mean tangential velocity along its trajectory (mm/s). The statistical distributions of each extracted feature are then analysed in detail, taking into account the two distinct experimental conditions: standard and metronome guided.

The complete outcome of this feature analysis for recurrent gestures is presented in the next chapter, among all of the results obtained with the movement analysis method described in this chapter.

Chapter 4

Results

In this chapter the results obtained for the two experiments, using the movement analysis procedure described in the previous chapter, are presented and analysed in detail.

4.1 First Experiment

4.1.1 Movement Recurrence and Regions of Interest

The proposed movement analysis procedure was conducted for the ten clarinet players in the first experiment. Two of them did not exhibit a significant movement recurrence pattern while performing the excerpt by Mozart. These two players are students, with less expertise than the others, and were discarded on the final step of the analysis. The other eight musicians exhibited high movement recurrence regions in the last three bars of Mozart's theme, similarly to the example shown in the previous chapter. Figure 4.1 shows the motion recurrence maps for these eight selected musicians.
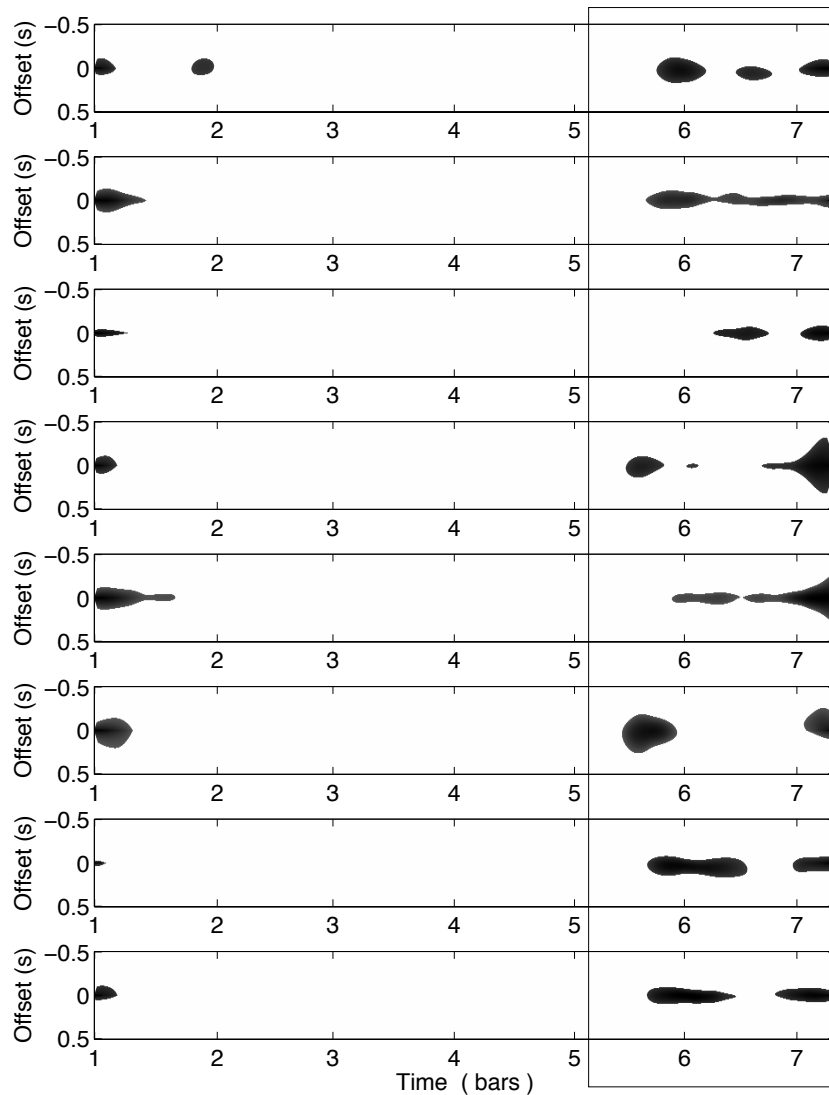
Figure 4.1: Clarinet bell motion recurrence maps for the performances of Mozart's excerpt by the eight selected musicians, with the last three bars highlighted. Dark regions indicate high recurrence.

In order to analyse the general pattern of recurrence for this group of players a global recurrence map was computed for Mozart's excerpt, by summing the maps of the eight selected musicians. The resulting map was then normalised and a new threshold of 0.4 was applied in order to highlight the global recurrences, that tend to be less prominent than in the case of an individual musician. The result is shown in Figure 4.2.
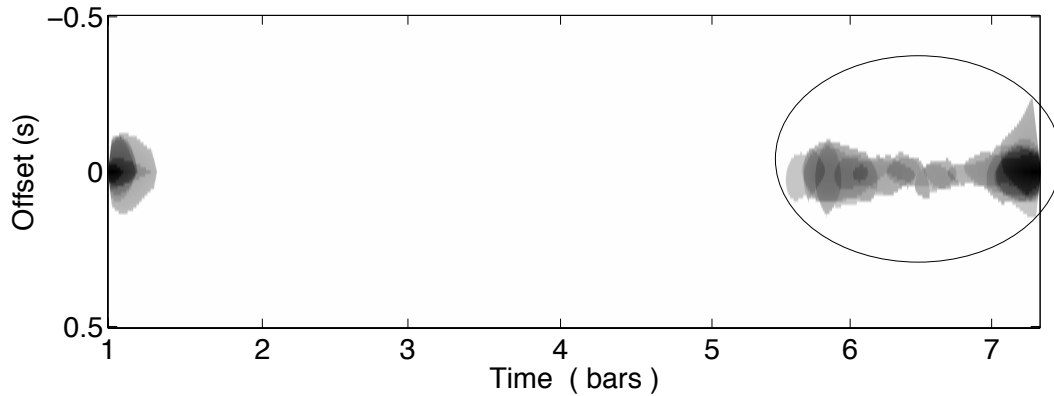
Figure 4.2: Clarinet bell global motion recurrence map for the performances of Mozart's excerpt by the eight selected musicians, with the last three bars highlighted. Dark regions indicate high recurrence.

Figures 4.1 and 4.2 clearly show that the main regions of interest occurred at the last three bars of Mozart's theme for all musicians. This region of the score, shown in Figure 4.3, is highlighted in the map figures, indicating that the recurrence pattern is very consistent across players and strongly related to the music structure of the excerpt.

Figure 4.3: Last three bars of Mozart's excerpt.

Inside these regions of interest for each clarinetist, the movement segments were then grouped into highly recurrent sequences of gestures (as exemplified in figures 3.7, 3.8 and in Appendix A), according to the procedure described in Section 3.5. The tangential velocity based movement segmentation and recurrence analysis revealed representative music related gestures, from which gestural features can be extracted for comparison. The occurrence of such recurrent gestures and their relation to the musical structure, specially for the most skilled players, constitutes strong evidence of the musical significance in the physical movements of musicians.

33

4.1.2 Gestural Features

The next step is to extract and analyse gestural features of this group of musicians, according to the definitions in Section 3.6. First, the recurrent gestures were subjected to Principal Component Analysis, to investigate their dimensionality.
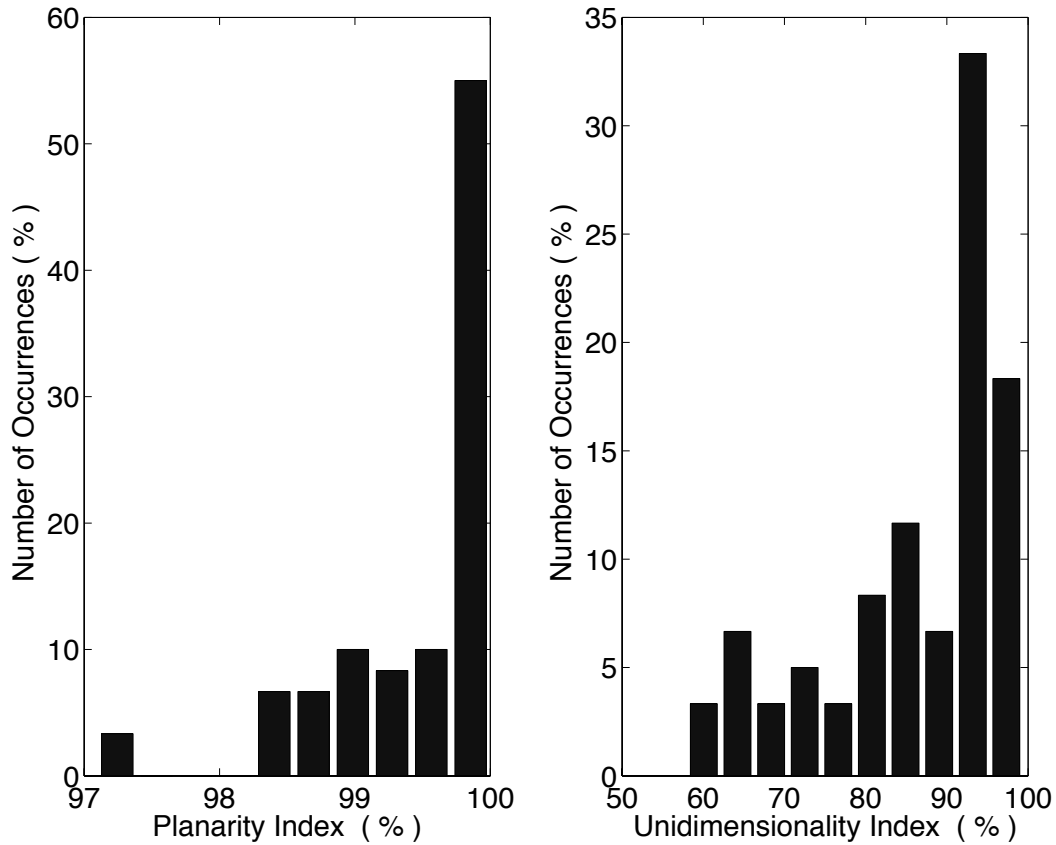
Figure 4.4: Histograms for the planarity and unidimensionality indexes of the trajectories of the clarinet bell gestures defined for the eight selected players in the first experiment. Total of 60 gestures.

The results of this analysis reveal that the trajectories of all recurrent clarinet bell 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 unidimensional, since 50% of them exhibit unidimensionality indexes above 95%, and 80% exhibit unidimensionality indexes above 80%. This

is illustrated in Figure 4.4, that shows histograms for the distributions of the planarity and unidimensionality indexes in this group of gestures.

These results indicate that the clarinet bell gestures executed during the musical performances have a very strong tendency to occur along a given plane, and in many cases also along a straight line in space. Since the clarinet bell movement is taken relative to a static reference point and the musicians performed standing up, gestural planarity cannot be assumed a priori and it thus corroborates previous assumptions and the procedure used to define recurrent gestures. This is a significant statement regarding the nature of these gestures and it will be further discussed in the next chapter.

The two experimental conditions were then compared regarding gestural prominence, based on the statistical distributions of three selected gestural features: the total spatial distance covered along its trajectory (mm), the time duration of the gesture (s), and the mean tangential velocity along its trajectory (mm/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, but with similar time durations. This is illustrated in Figure 4.5, that shows box plots for the statistical distributions of these three features.

Paired t-tests were performed over the difference between the two experimental conditions, relative to these three features. The null hypothesis was rejected at the 1% significance level ($\alpha = 0.01$) in the cases of gestural amplitude (p value < 0.001) and mean velocity (p value < 0.001), confirming the statistically significant distinction between the experimental conditions regarding these features. In the case of gestural duration this test indicated no statistically significant distinction between the experimental conditions.
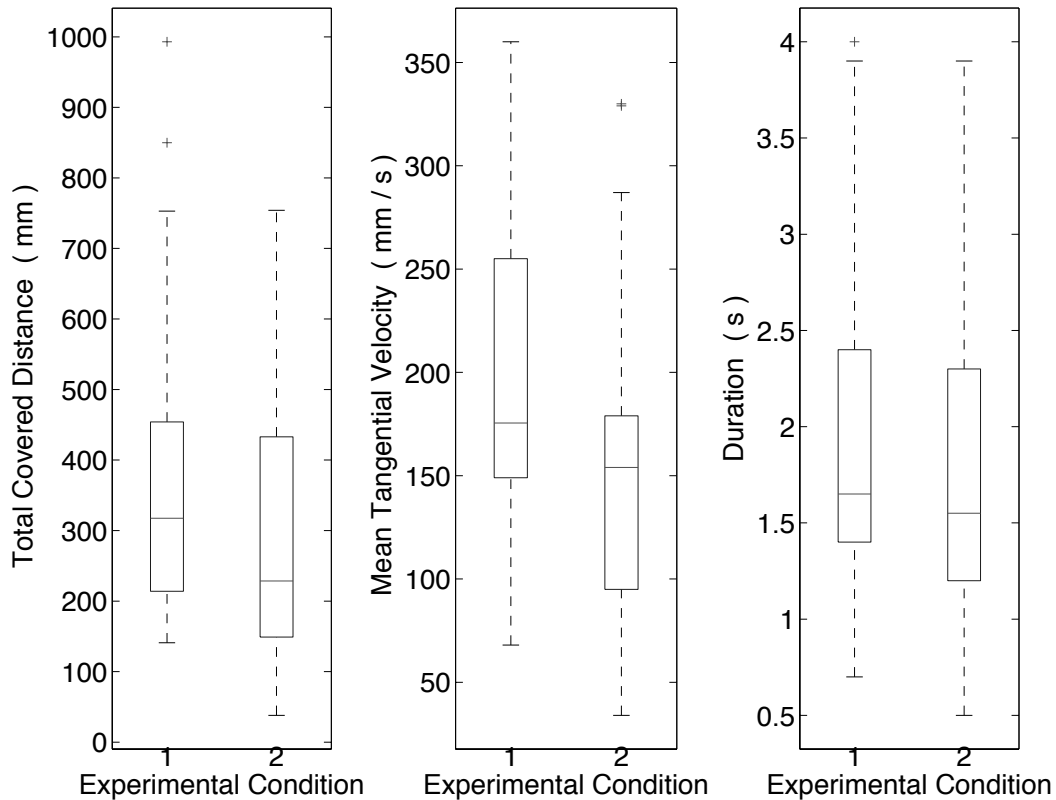
Figure 4.5: Box plots for the statistical distributions of three gestural features of the clarinet bell gestures defined for the eight selected players in the first experiment, comparing the two experimental conditions. 1: standard ; 2: metronome. Central line: median ; Edges: 25th and 75th percentiles. Total of 60 gestures.

This means the metronome had a clear restraining effect on the gestural extension and velocity within the regions of interest, but did not change the location or duration of the clarinet gestures. This result reinforces the hypothesis of a musical significance in the musicians' ancillary movements, related to their expressive intentions and to local musical targets (Vines, Dalca, and Wanderley, 2006a).

4.1.3 Audio Analysis

The main regions of movement recurrence were observed at the last three bars of Mozart's theme (shown in Figure 4.3), just where there is more room for expressiveness, due to the greater rhythmic and melodic variation and also to the

conclusion of the musical phrase. This is coherent with the idea of a strong expressive content being imposed by the musicians in this final part of the excerpt (using small deviations of note duration, articulation, intensity, pitch and timbre), and being reflected in their corresponding physical gestures. In order to further investigate this, the relative note durations were analysed along the performances, since note duration manipulation is one of the most noticeable and consolidated indicators of expressive intentions by the music performers (Gabrielsson, 1995; Juslin, 2000; De Poli et al., 2004).

This is done through the calculation of the inter-onset intervals (IOI's), based on the note onsets extracted from the audio signals. In order to get a relative idea of the note durations, according to their canonical durations in the score, the IOI's are normalised 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 horizontal line indicates a note duration manipulation by the musician, with upward deviations representing a relatively longer musical note and downward deviations representing a relatively shorter musical note.

Figure 4.6 exemplifies this analysis for two musicians. Black curves represent the standard performances, grey curves represent the metronome performances, grey vertical bars show the relative nominal value of the notes in the score, used in the normalisation, and the black horizontal dashed line indicates the canonical values. The last three bars of the excerpt are highlighted in the rectangle.

The evolution of these note inter-onset intervals also reveals a great contrast between the initial and final part of the excerpt, as illustrated by Figure 4.6. In the first four bars there is little manipulation of the note durations by the musicians, while in the last four bars (inside the rectangle) the musicians execute significant and recurrent manipulation over the note durations, as indicated by the large
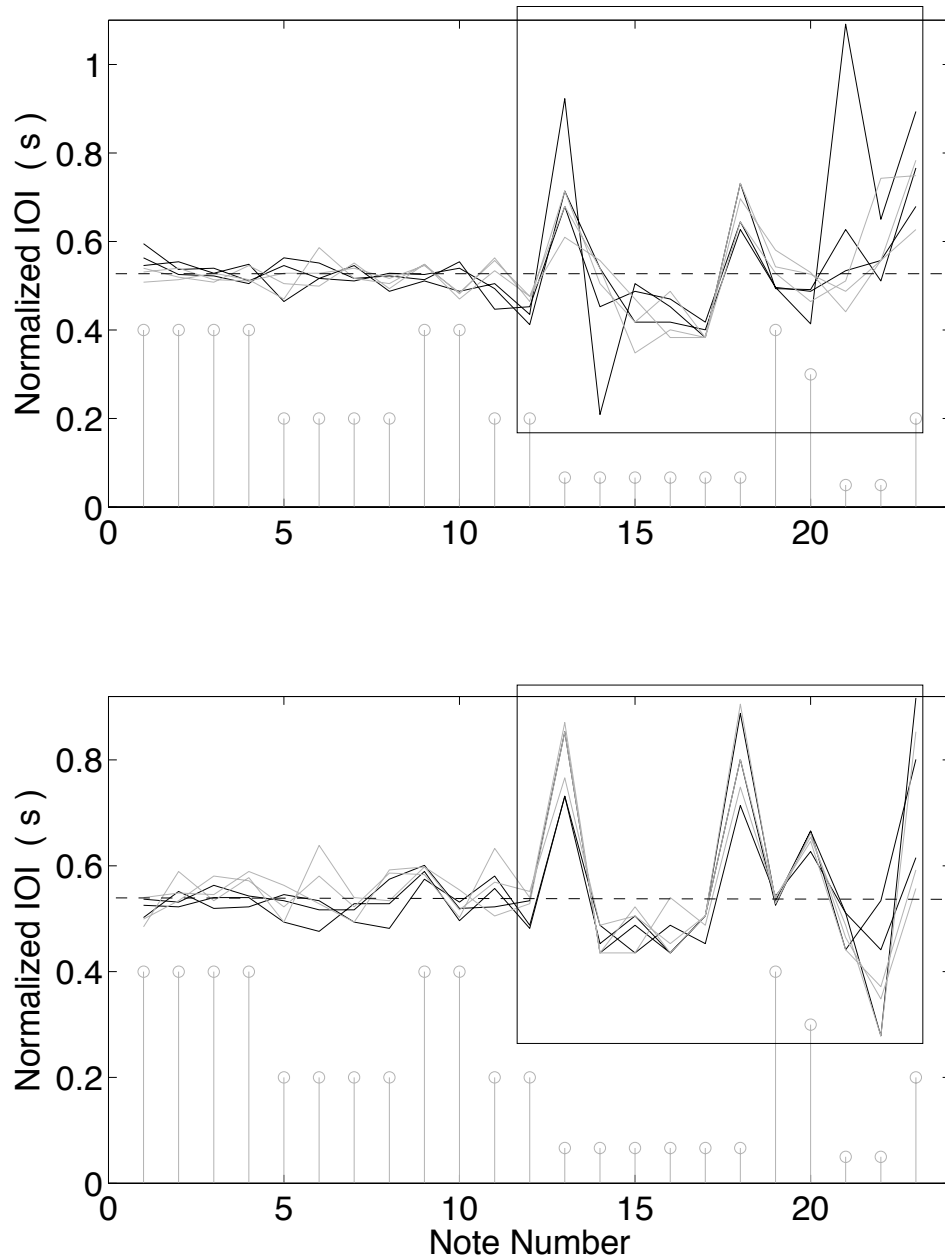
Figure 4.6: Note inter-onset intervals for the performances of Mozart's excerpt by musicians 1 (top) and 7 (bottom), normalised relative to a quarter-note. Standard performances in black, metronome performances in grey, relative nominal score value of the notes in the grey vertical bars. The horizontal dashed line shows the canonical values and the rectangle highlights the last three bars.

deviation of the IOI curves around the horizontal dashed line. This was observed in the performances of almost all of the musicians in the first experiment, and is in accordance with the previous assumption that a strong expressive content is imposed by the clarinetists in the last three bars of this Mozart's theme.

The results of this audio analysis thus corroborate the hypothesis that the strong expressive intentions present in the final part of these performances are reflected in the physical gestures of the clarinetists, since both movement and note duration analysis led to highly related conclusions. This will be further discussed in the text, also taking into account the results of the second experiment.

4.2 Second Experiment

4.2.1 Movement Recurrence and Regions of Interest

The movement analysis procedure described in the previous chapter was also conducted for the 13 clarinet players who performed the excerpt by Brahms in the second experiment. The two students discarded from the first experiment also did not show a significant recurrence pattern in this case and were again discarded from the final analysis. Out of the remaining 11 musicians, another one was discarded. He was clearly uncomfortable with the experiment and also did not show any recurrent patterns.

Ten players were then used in the final step of the analysis. Figure 4.7 shows the motion recurrence maps for the selected musicians and Figure 4.8 shows the global recurrence map for the Brahms' excerpt, computed by summing the maps of these ten clarinetists, as explained in the previous section.

The inspection of figures 4.7 and 4.8 reveals that the main regions of interest for these musicians occurred mostly around four common regions in Brahms' excerpt.
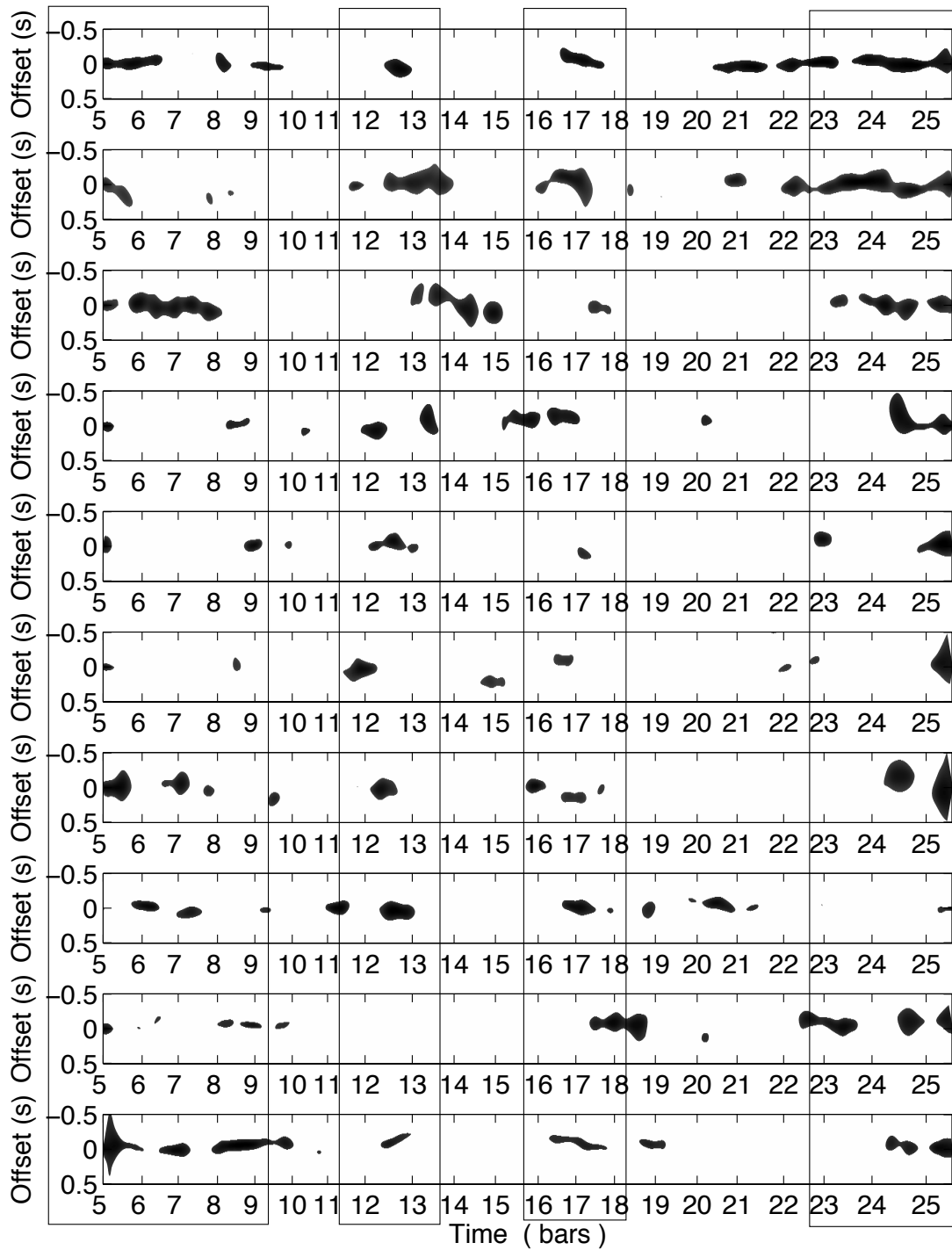
Figure 4.7: Clarinet bell motion recurrence maps for the performances of Brahms' excerpt by the ten selected musicians, with four musical regions highlighted. Dark regions indicate high recurrence.
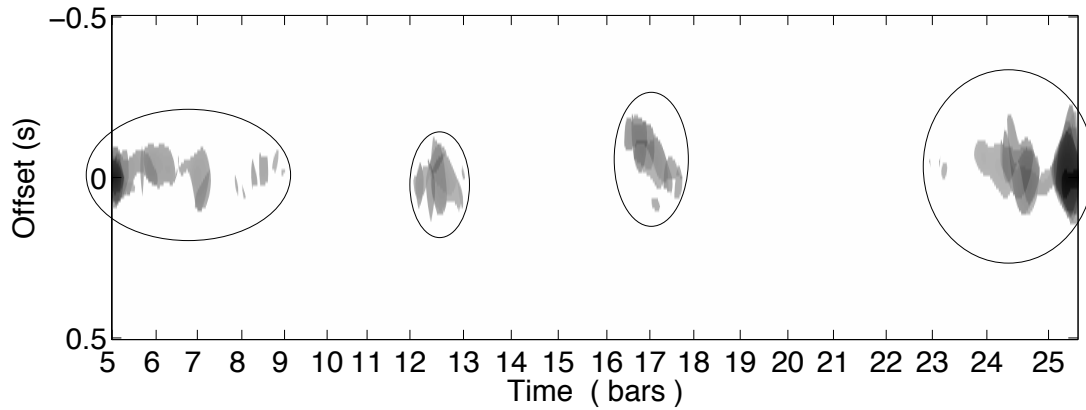
Figure 4.8: Clarinet bell global motion recurrence map for the performances of Brahms' excerpt by the ten selected musicians, with four musical regions highlighted. Dark regions indicate high recurrence.

Figure 4.9: Excerpt from Brahms' Clarinet Sonata No.1. The four regions highlighted in the recurrence maps are indicated in the score.

These regions are indicated on these figures and also on the musical score in Figure 4.9. The most prominent and recurrent clarinet gestures were executed in these regions, and the percentage of occurrence of such gestures inside them is very high. For this group of musicians on average, recurrent clarinet gestures were observed inside each of these regions in 85% of the cases, much more often than in other parts of the excerpt. This general behaviour indicates a strong relation between the musical structure of the excerpt and the occurrence of recurrent gestures in the performances by these musicians. These results also suggest that these regions

41

of the excerpt correspond to notable musical targets for these musicians while playing. This will be further discussed in Section 5.5.

Inside the regions of interest for each clarinetist, the movement segments were grouped into recurrent sequences of gestures, according to the procedure described in Section 3.5. As an example, Figure 4.10 illustrates the definition of the regions of interest and their constituent gestures for the three performances of Brahms' excerpt by Musician 1 and Figure 4.11 shows the 2D trajectories of the gestures (xy plane, as shown in Figure 2.3) inside the first region of interest for this player. Additional examples are presented in Appendix A.
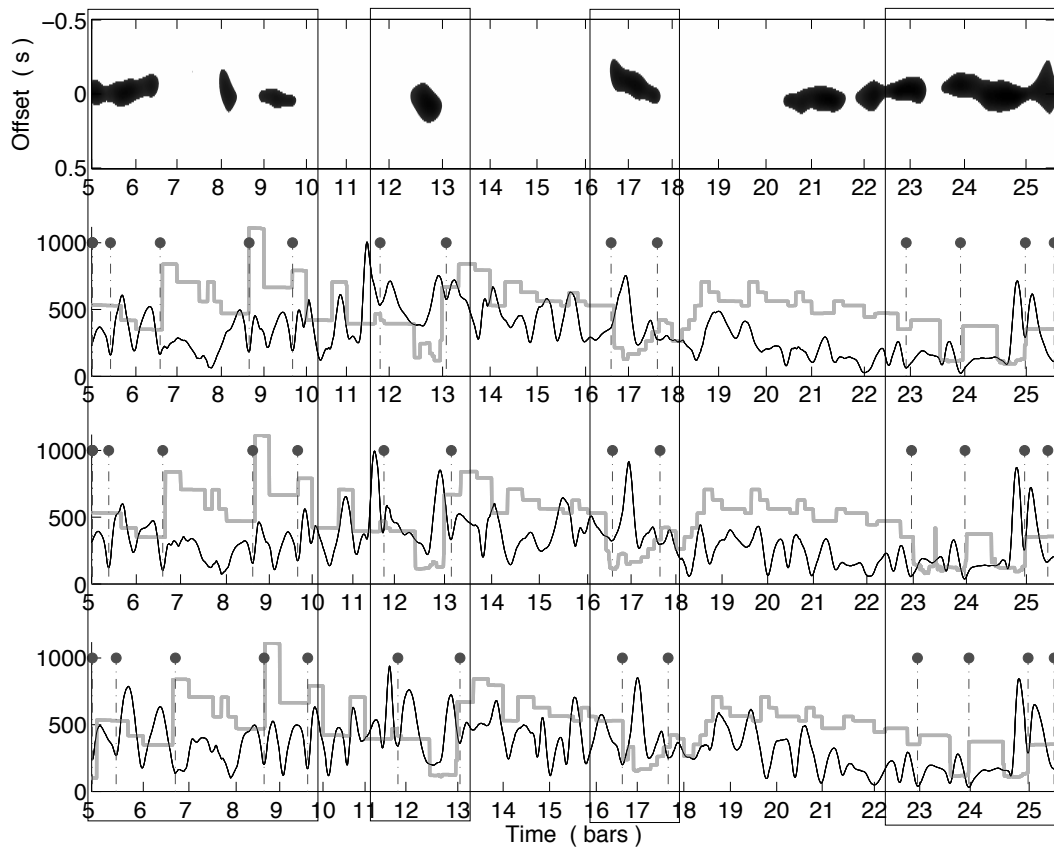
Figure 4.10: Regions of interest for the three performances of Brahms' excerpt by Musician 1, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map, with dark regions indicating high recurrence.

As in the case of the first experiment, and according to figures 4.10 and 4.11, the number, sequence, location, geometry and duration of these gestures are all highly recurrent within each region of interest. The occurrence of such highly recurrent gestures for all players, in close relation to the music structure of a different excerpt, constitutes additional evidence of the musical significance of these gestures, corroborating the previous findings in this study.
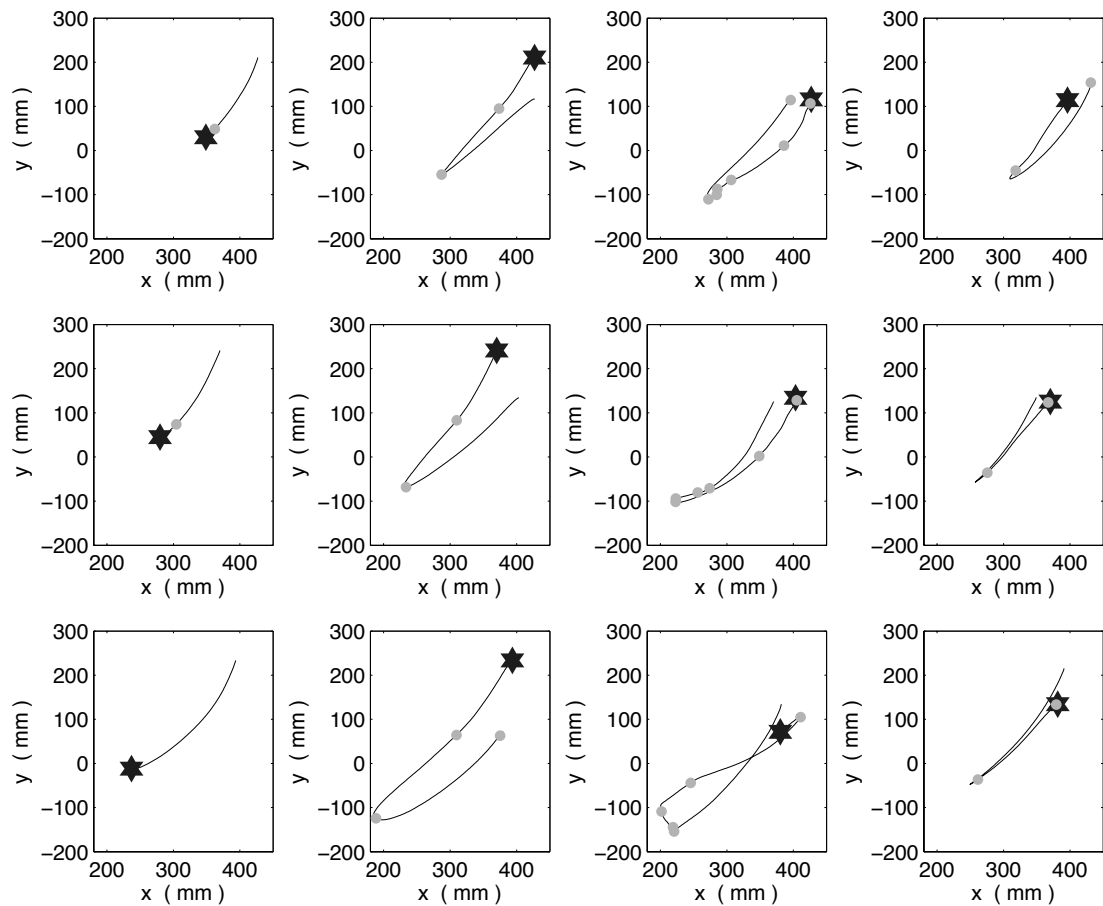
Figure 4.11: 2D trajectories of the four gestures inside the first region of interest for the three performances of Brahms' excerpt by Musician 1. The star marks the initial point and the circles indicate the note onsets.

4.2.2 Gestural Features

With the definition of recurrent clarinet bell gestures in each region of interest, it is possible to perform the feature analysis of gestures proposed in Section 3.6. In the case of the second experiment there are only standard performances of the excerpt by Brahms, so the recurrent gestures were only subjected to Principal Component Analysis, in order to investigate the dimensionality of their trajectories.
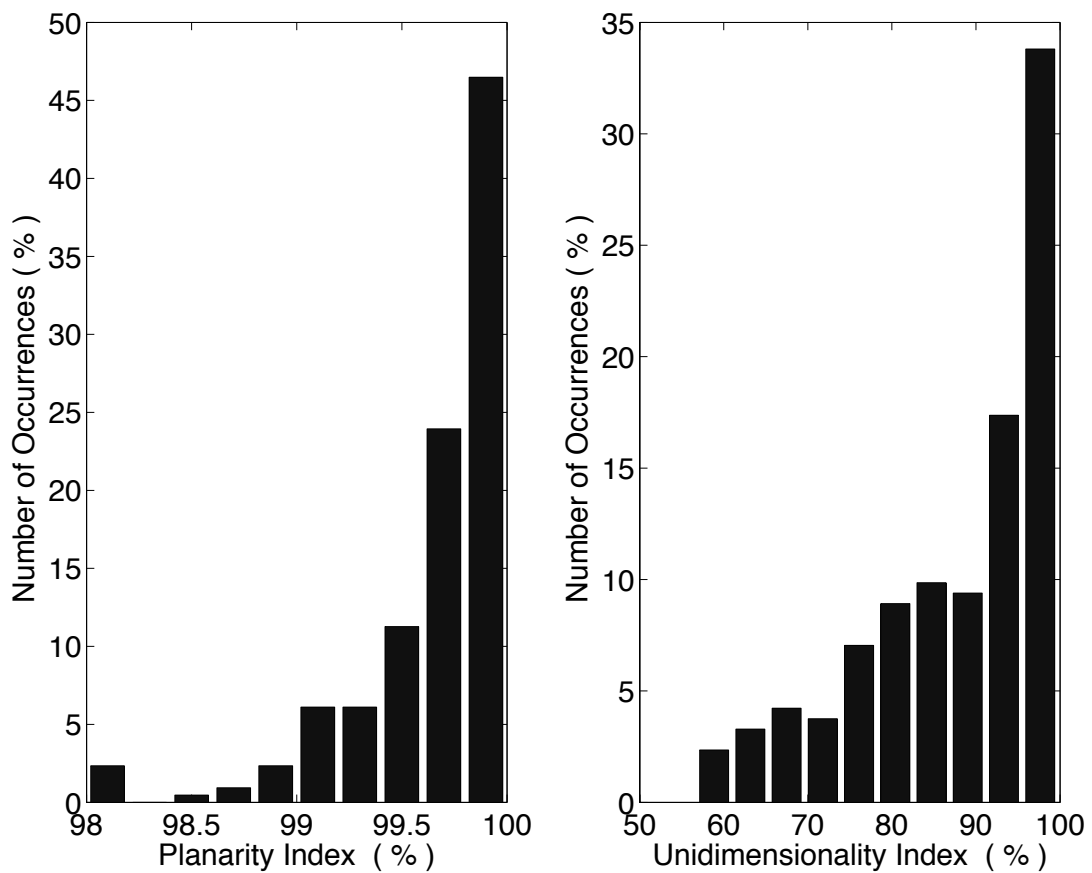
Figure 4.12: Histograms for the planarity and unidimensionality indexes of the trajectories of the clarinet bell gestures defined for the ten selected players in the second experiment. Total of 213 gestures.

The results of this analysis for this group of gestures reveals that all of their trajectories are highly planar, with planarity indexes always above 98%. About half of them also exhibit unidimensionality indexes above 95%. This is illustrated

in Figure 4.12, that shows histograms for the distributions of the planarity and unidimensionality indexes in this group of gestures.

These results are in accordance with the ones obtained in the first experiment, and indicate that these clarinet bell gestures have a very strong tendency to occur along a given plane, and in about half the cases also along a straight line in space. The implications of this will be discussed in detail in the next chapter.

4.2.3 Audio Analysis

According to the results presented in this section, the most prominent and recurrent clarinet gestures were executed by the musicians around four regions in the excerpt by Brahms. In these regions recurrent sequences of gestures were found much more often than in the rest of the excerpt in general.

Similarly to what was done in the first experiment, an audio analysis was conducted over this data set, in order to investigate the expressive content of the performances through the examination of note duration manipulations by the players. The goal is to relate the results of this audio analysis to these patterns of movement recurrence.

The excerpt by Brahms selected for the second experiment is longer and much more complex than the theme by Mozart selected for the first experiment. This makes it very difficult to effectively analyse the evolution of the duration of all its individual notes. Because of this, the analysis in this case investigates the manipulation of bar durations by the musicians.

This is done according to the following procedure. For each performance the duration of every bar is calculated by summing the duration of its constituent notes. These bar durations are then normalised using the mean value of bar duration in that performance. This way it is possible to get a relative idea of

bar duration manipulation by the player, in relation to the mean value of each performance, and also to compare it between various performances, regardless of the tempo employed in each individual case.

The goal of this analysis is to obtain a general measure of bar duration manipulation for the ten selected players, that can be related to their expressive intentions towards this excerpt. This can be done through the calculation of the standard deviation of the normalised bar durations. For this, the standard deviation of the normalised duration of each bar, over the 30 performances by this group of players, is computed and evaluated.

The assumption behind this analysis is that the bars in the excerpt with weaker expressive content will lead in general to small duration manipulations by the players, resulting in normalised bar duration values close to unity and consequently to small values of standard deviation for this feature. On the other hand, bars with stronger expressive content will lead to substantial duration manipulations overall, resulting in scattered values of normalised duration and thus to large values of its standard deviation.

Figure 4.13 illustrates the overall results of this audio analysis. It presents the standard deviation values, given in percentage points, for the normalised durations of each bar in Brahms' excerpt, over the 30 performances evaluated in this experiment.

As can be seen in Figure 4.13, the four most prominent peaks in the curve representing the evolution of standard deviation of normalised bar durations occurred at bars 8, 12, 16 and 24. This means that the bars in the excerpt that led the players to the most significant duration manipulations are all located inside the four main regions of movement recurrence in the music, indicated in figures 4.7, 4.8 and 4.9.
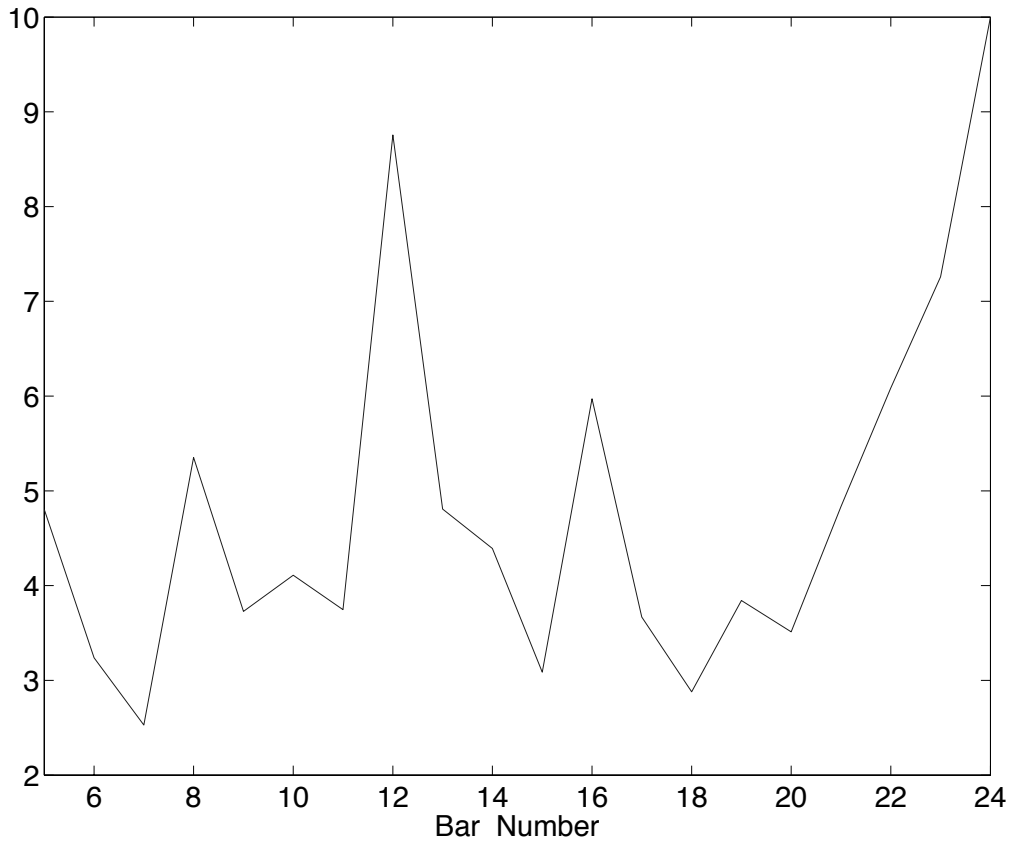
Figure 4.13: Standard deviation (%) for the normalised durations of each bar in Brahms' excerpt, over the 30 performances in the second experiment.

The fact that the four regions in the excerpt where the most consistent movement recurrence patterns occurred for this group of clarinetists are so highly related to musical moments in the performances with apparent strong expressive content is another very strong result. It provides additional evidence that the ancillary physical gestures employed by musicians during performances are closely related to their musical intentions towards the musical realisation.

Like in the first experiment, the results of this audio analysis also corroborate the hypothesis that the strong expressive intentions present in these regions of the excerpt are reflected in the physical gestures of the clarinetists, since both movement and bar duration analysis led once more to highly related conclusions. This will be further discussed in the next chapter, more specifically in Section 5.5.

Chapter 5

Discussion

5.1 Movement Analysis Method

A fundamental step in the motion analysis of clarinet players is to find an effective scalar representation of their movements in time and space. By transforming the tridimensional motion data into a unidimensional temporal signal, similar to the ones describing the respective acoustic data, it is possible to use signal processing techniques to analyse both aspects of the musical performances in conjunction. This was done through the estimation of the tangential velocity of the clarinet bell.

The tangential velocity representation of movement provides a simple way to visualise and analyse the temporal evolution of the movements made by the musicians, according to the musical structure of the excerpt being performed (as shown in Figure 3.2). It also makes it possible to segment these movements into short and basic units (as shown in Figure 3.3), used afterwards to define gestures and to conduct a local parametrical analysis.

The results of the study indicate that this motion representation is indeed very efficient for the intended movement analysis. Namely, the high recurrence

in the tangential velocity curves inside the defined regions of interest, along all the performances by each musician (as seen in Figure 3.4), shows that this is an effective representation of the player's movements, retaining and unveiling its invariant aspects in a simpler unidimensional form. Many times the corresponding movements in successive performances by a musician are very similar but not identical, constituting different ways to convey the same musical intention. Even in these cases the tangential velocity can unveil the high levels of recurrence.

The use of tangential velocity movement segmentation was also the basis for the definition of highly representative and recurrent physical gestures, with complete and primitive geometrical trajectories, formed by the grouping of such movement segments (as seen in figures 3.7, 3.8, 4.10, 4.11 and in Appendix A). These physical gestures were unveiled with the aid of a movement recurrence analysis, also based on this motion representation, using time-warped versions of the tangential velocity curves as temporal signals fed to an instantaneous correlation algorithm.

These conclusions show that the tangential velocity of the clarinet bell is in fact a good choice for motion representation during musical performances, and that its local minima correspond indeed to significant inflection points in the players' movements, as previously stated.

5.1.1 Related Work

The gestures were extracted based on a single low-level kinematic feature, the tangential velocity of the clarinet bell. Higher-level features were already used in previous studies by Camurri et al. (2004) and Desmet et al. (2012), defining gestures based on Laban's effort space, or a combination of several markers through Principal Component Analysis, among other methods.

The choice of method in this study was guided by simplicity and justified by

49

the assumption that the clarinet bell is a good general indicator of movement for the player, providing a single trajectory from which geometrical and temporal features can be extracted and locally related to the score. Taking the clarinet bell motion relative to a static reference point was also used as a simple way to combine the movements of several areas in the body into one analysis point. Even so, several other markers were used in the experiments and may be used in the future to analyse different body movements through high-level features, adding to the current results.

5.2   Movement Recurrence

In order to establish relations between the musicians' movements, their expressive intentions and the musical structure of the excerpts, it is important to find out which movements occur recurrently at a given point in the music, over all the performances by each clarinet player. As previously stated, it can be expected that many aspects of music performance remain mostly unchanged over different executions by an expert player, including the musical intentions and the movement patterns. This was investigated with the use of the motion recurrence map analysis described in Section 3.4.

This analysis shows that most of the musicians in the study execute highly recurrent movements at specific moments along the excerpts. Apart from these moments, their movements tend to occur according to a more varying pattern, in terms of spatial evolution and temporal location inside the music. This suggests the existence of key moments in the excerpts, where the recurrent movements employed by the musicians seem to indicate clearer expressive intentions inside the performances, related to specific musical passages. These music related considerations will be discussed further in Section 5.5.

These results indicate that the physical movements employed by these clarinet players are somehow important for their intended performances, given their high recurrence and strong relation to the musical structure, at key moments in the two excerpts. Even so, it is not possible to state that these movements are strictly fundamental for such performances, since a few musicians do not seem to employ significant recurrent movements while performing the excerpts, even if they are the minority in the group (around 20% of them) and also the least experienced ones.

For each musician, regions of interest were then defined around the high movement recurrence areas in the excerpts, in order to look for recurrent physical gestures. Inside these regions of interest, the location, number, sequence, geometry and duration of these gestures are all highly recurrent. The definition of high movement recurrence regions of interest in the performances can thus unveil representative music related physical gestures. The strong recurrence and relation to the musical structure found in these physical gestures, specially for the most skilled players, corroborate the hypothesis of a musical meaning behind the ancillary movements of the musicians, related to their expressive intentions and important for the desired musical outcome.

5.2.1 Related Work

Other methods have also been used to measure recurrence in studies related to music performance. Varni, Volpe, and Camurri (2010) used Recurrence Quantification Analysis (RQA) to analyse synchronisation and interaction in performances by a violin duo. This method is used to compute recurrence plots based on the phase synchronisation within a signal or between signals, using the phase space trajectory of complex dynamical systems that model the players.

The method developed in this study to investigate motion recurrence was chosen because it is very straightforward and based on a simple measure of correlation. The resulting recurrence plots can be easily displayed in alignment with the music structure (through the pitch curve and note onsets) and the velocity signals representing the movements, preserving the original time axis. Several recurrence plots can also be summed directly to analyse the global recurrence over different groups of performances and musicians. The temporal fluctuation in the location of the recurrent movements inside the excerpt over different performances can also be assessed, through the time offset parameter represented on the vertical axis of the maps.

5.3 Gestural Dimensionality

With the definition of recurrent clarinet bell gestures inside the regions of interest in each performance, a local parametrical analysis was conducted over their spatial trajectories. In order to analyse their dimensionality, the gestures had their trajectories subjected to a principal component analysis, defining their planarity and unidimensionality indexes in each case.

The results for the two experiments are presented in figures 4.4 and 4.12. They show that all of these gestures are highly planar and about half of them are also highly unidimensional. This means that the clarinet bell gestures executed by the musicians in this study have a very strong tendency to occur along a given plane in space, and also along a straight line in many occasions. The clarinet bell movement is taken relative to a static reference point and the musicians performed standing up, so this gestural planarity cannot be assumed a priori, corroborating then the procedure used to define recurrent gestures.

These results are in accordance with the physical characteristics and constraints regarding the experiments. Based on the principle of conservation of angular momentum, circular movement trajectories tend to stay inside a principal plane, or around a fixed rotation axis, since an additional force would be required to shift this rotation axis. Also, one end of the clarinet is physically coupled to the mouth of the musician, which makes it difficult to execute complex and fully tridimensional movements using the instrument.

The results are also in accordance with conclusions obtained in previous related studies (Teixeira et al., 2010; Teixeira, Loureiro, Yehia, and Wanderley, 2013; Teixeira, Loureiro, Wanderley, and Yehia, 2014), that have already shown this tendency of occurrence of planar physical gestures during clarinet performances, also indicating that the trajectories of significant gestures tend to occur around an equilibrium position, related to the centre of gravity of the musician, with partially elliptical round-trip paths and a clear return point, according to theories by McNeill (2007). The new results in this study corroborate these conclusions regarding the spatial nature of such gestures, which are now confirmed for a large set of clarinet players and physical gestures. Besides validating all previous conclusions about the spatial trajectories of these gestures, for a much wider data set, the results presented in the current study also indicate, as a new conclusion, the occurrence of many unidimensional physical gestures during the performances.

According to these conclusions, these recurrent clarinet bell gestures can be effectively represented in space by its two principal components, and in about half the cases by only one principal component. This assumption can significantly simplify the visualisation and local parametrical analysis of such gestures, by reducing the dimensionality of their representation.

5.4 Experimental Conditions

In the first experiment the clarinetists were asked to play according to two experimental conditions, standard and metronome guided, in order to quantitatively compare between free musical performances and performances restrained by the metronome beat. These two conditions were compared regarding gestural prominence, based on the statistical distributions of three gestural features: total distance covered, mean tangential velocity and time duration (as seen in Figure 4.5). According to the results, the metronome had a clear restraining effect on the gestural extension and velocity within the regions of interest, but did not change the location or duration of the clarinet bell gestures.

This statistical analysis provides a general idea about the influence of the metronome over all of the performances, but each musician is influenced by the metronome guiding in a particular way. Some musicians seem to be only slightly influenced by the metronome, retaining their general movement patterns, with an occasional small decrease in gestural amplitude and velocity. Other musicians, while retaining most of their basic movement patterns, exhibit a strong decrease in the amplitude and velocity of their recurrent gestures.

An interesting example of this sort of behaviour is shown in Figure 5.1. This figure shows the most prominent recurrent gesture for Musician 8, in the region of interest at the end of Mozart's excerpt. It is possible to see that this musician executes a geometrically similar gesture at this musical location in all of his 6 performances of the excerpt, but the amplitude of this gesture is much smaller during the metronome guided performances, shown in the right column of the figure. This observation indicates that this musician executes a miniature version of the recurrent gesture under the influence of the metronome. This suggests that the metronome inhibits the movements of the musician, which become slower and

smaller, but do not eliminate them entirely, since a miniaturised gesture remains at the same musical location and with the same geometrical pattern.
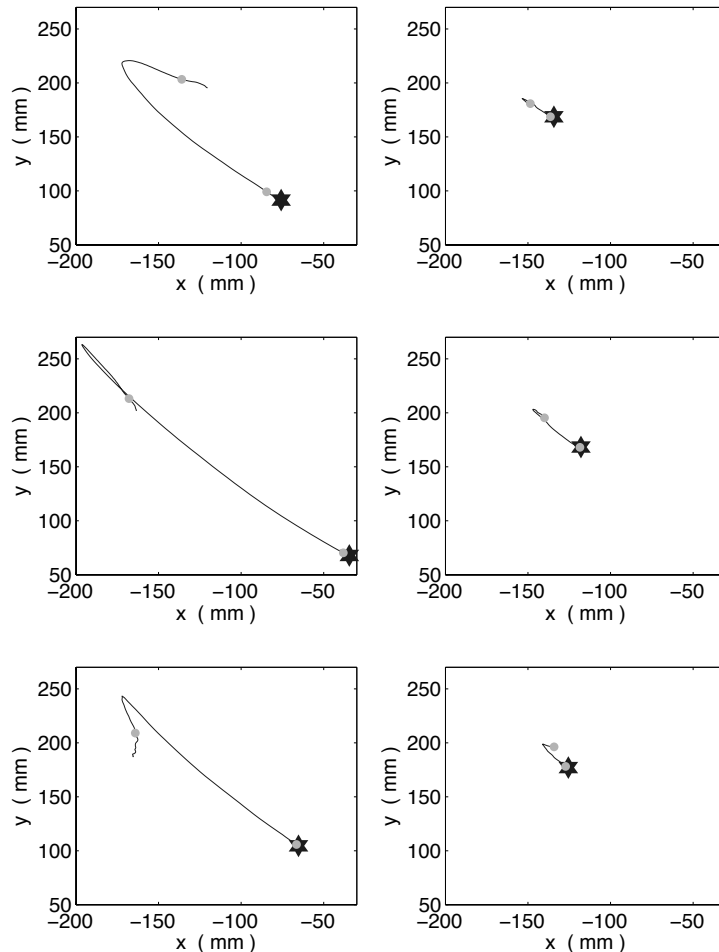

Figure 5.1: 2D trajectories for a gesture defined in the region of interest at the end of Mozart's excerpt, for the 6 performances by Musician 8. The star marks the initial point and the circles indicate the note onsets. Left column: standard performances, right column: metronome guided performances

Another interesting and quite stronger influence of the metronome over the performances was observed for a couple of musicians. An example of this sort of behaviour can be seen in Figure 5.2. This figure shows a motion recurrence map for the performances of Mozart's excerpt by Musician 2 (similar to Figure 3.4). In this case the motion recurrence map relative to all of her six performances shows
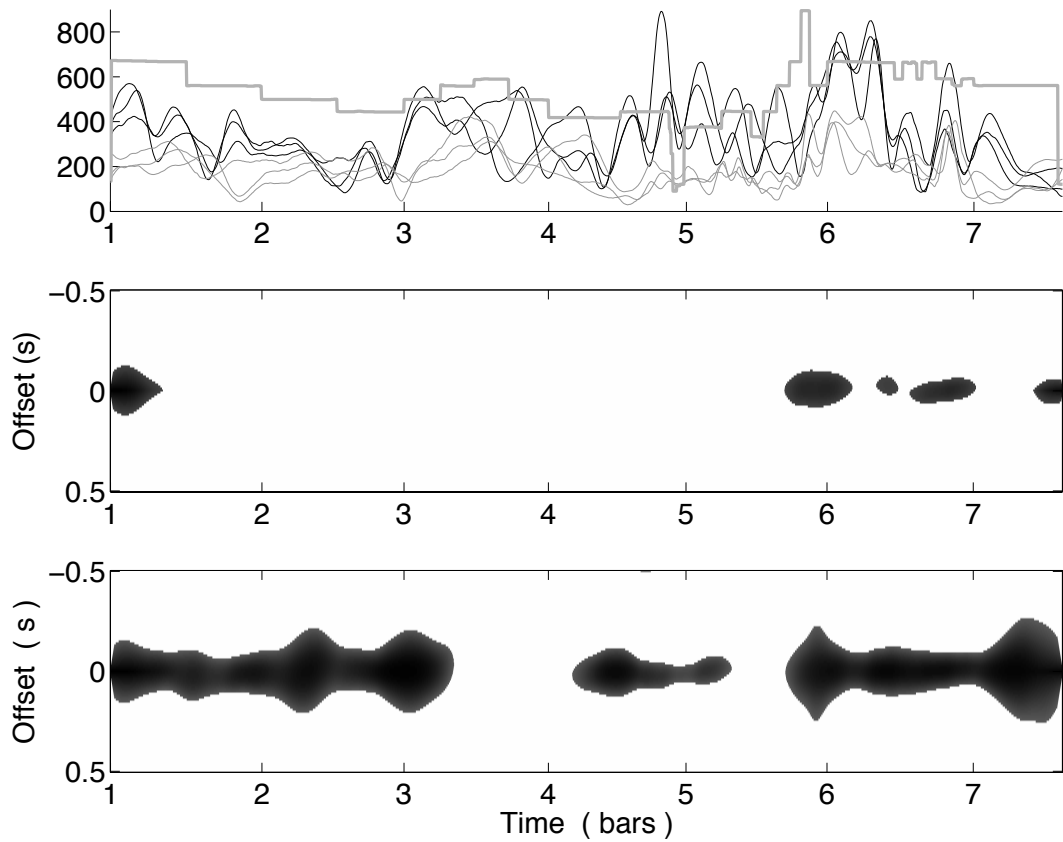
Figure 5.2: Clarinet bell motion recurrence map for Musician 2 over her six performances of Mozart's excerpt, and over her three standard performances only. Top plot: clarinet bell tangential velocity curves (mm/s), standard performances in black, metronome performances in grey, pitch (Hz) in the thick grey curve. Middle plot: motion recurrence map for all six performances. Bottom plot: motion recurrence map for the three standard performances only. Dark regions indicate high recurrence.

small regions of recurrence, as seen in the middle plot of the figure. But looking to the tangential velocity curves shown in the top plot of the figure, it is possible to notice that the three black curves, relative to the standard performances, seem to be highly correlated. Based on this observation, a motion recurrence map was calculated for the three standard performances only, as seen in the bottom plot of the same figure. Analysing this figure it is possible to see that this musician exhibits a highly recurrent movement pattern over her three standard performances of the

excerpt, and also that the use of the metronome totally altered this movement pattern, giving rise to a different kind of interpretation. This conclusion suggests that the two musicians that exhibited this sort of behaviour are strongly influenced by the presence of the metronome, executing a clear movement pattern, recurrent over their standard performances, that is totally disrupted by the metronome guiding.

These observations indicate that most of the studied musicians are slightly influenced by the metronome, since their recurrent gestural patterns are not disrupted by its presence, only suffering varying degrees of decrease in their amplitudes and tangential velocities. On the other hand, a smaller group of musicians seem to be strongly influenced by the metronome, that alters their movement patterns completely.

For this second group of musicians, the presence of a recurrent movement pattern in their standard performances, that does not occur during their metronome guided performances, is a strong evidence of the relation between such recurrent movements and their respective expressive intentions towards the music. With the metronome making it difficult for them to impose their musical intentions during the performance, their movement patterns would also be altered accordingly. For the first group, on the other hand, even if the metronome reduces their ability to impose expressive intentions during the performance, causing a consequent decrease in their gestural amplitude and velocity, the recurrent gestures still remain at their musical locations, indicating a clearer overall musical intention by these musicians, even when following the metronome. Again, the stronger movement recurrence observed in this first group, together with the decrease in gestural amplitude and velocity imposed by the presence of the metronome, are still good evidences of the relation between these gestures and the expressive content of the performances.

All of the conclusions discussed in this section corroborate the hypothesis of a musical significance in the musicians' ancillary movements, related to their expressive intentions and to local musical targets in the performance of the excerpt.

5.5   Musical Context

5.5.1   Mozart's Excerpt

The results of the movement analysis indicated that the most significant recurrent clarinet bell gestures occurred during the final part of Mozart's excerpt (bars 5-7), shown in Figure 4.3. The analysis of note inter-onset intervals also revealed a large contrast between the two regions in the excerpt, with this final part containing the most significant and recurrent note duration manipulations.

If compared to the initial part of the excerpt, this final passage exhibits a wider pitch range, a larger variation of note durations and also contains the fastest notes. It consists of a long note, preceded by triplets, and followed by a trill. This creates a strong contrast with the smooth sequence of longer notes at the beginning of the excerpt. The fast triplets phrase is, in this sense, a key transition moment in this musical theme, separating these two distinct musical parts.

This might explain why all musicians execute their most significant recurrent clarinet gestures around this fast triplets phrase. Some of the musicians started these gestures at the beginning of this musical phrase, while others started them at the end of this phrase. Even so, all of the analysed recurrent gestures occurred in the last three bars of the excerpt, and seem to have a direct relation with this fast musical phrase. These clarinet gestures can be seen as punctuation movements at the end of this distinctive musical phrase, or even as global finalising gestures for the excerpt itself, especially in the cases where they start at the end of this phrase.

Some musicians start these gestures at the beginning of this phrase instead, and this could also be to help them execute it according to their musical intentions, altogether with the ending of the excerpt. This might indicate that some musicians impose their finalising musical intentions from the beginning of the triplets phrase, while others do so only after this phrase.

To sum up, both clarinet bell movement and note duration analysis led to related results, indicating that the most significant recurrence patterns occur in this region at the end of Mozart's excerpt, especially around the fast triplets musical phrase, which seems to be indeed a key transition point in the music. This corroborates the hypothesis of a strong expressive content being imposed by the musicians in this final part of the excerpt, and being reflected in their corresponding physical gestures. These conclusions are also in accordance with the musical structure of the excerpt, as discussed in this section.

5.5.2   Brahms' Excerpt

According to the results presented in Section 4.2, the most prominent and recurrent clarinet bell gestures were executed by the musicians around four regions in Brahms' excerpt (indicated in Figure 4.9). The corresponding audio analysis also showed that the bars in the excerpt that led the players to the most significant duration manipulations are all located inside these same regions in the music. This excerpt is longer than the one used in the first experiment, with more complex harmonic and melodic content, providing more room for expressiveness. This results in the occurrence of a more complex and extensive movement recurrence pattern, centred mainly around these four musical regions.

In order to better understand the relation between the musical structure of the excerpt and the occurrence of this gestural and expressive pattern, the melodic and

harmonic content of the excerpt is examined in this subsection. For that, Figure

5.3 shows the full score of the excerpt with these four regions highlighted.

Figure 5.3: Excerpt from Brahms' Clarinet Sonata No.1, Op.120 No.1 with the
piano. The 4 regions highlighted in figures 4.7 and 4.8 are indicated in the score.

The first of these regions in the score concludes at bar 8, as seen in Figure 5.3,

where a Neapolitan sixth subdominant chord prepares the move to the dominant

region, reached at bar 12. That peculiar harmony, together with the wide upward

interval leap in legato on the clarinet at this bar (minor tenth), might explain the

recurrence observed in this region. It is worth to note that this minor tenth is the

wider of a sequence of upward interval leaps on the clarinet melody started at bar 6, a melodic recurrence upon which the main musical idea of the whole sonata is based.

The second of these regions in the score is centred around bar 12, where this melodic sequence finally concludes into the dominant region, resulting in a strong feeling of punctuation and harmonic tension that makes this bar a notable musical target.

The third of these regions in the score is centred around bar 16, where the dominant is restated in the three bars after bar 12, with the use of melodic variations of the main theme. This might be seen as a recurrence of the dominant move in bar 12, resulting in a second musical target with more accentuated expressive tension.

The fourth of these regions in the score corresponds to the last three bars of the excerpt (23-25), which prepares the return to the Tonic at bar 25, in a diminuendo started at bar 21, evoking a strong feeling of harmonic resolution. However, this relaxation on the tonic is reached in sudden *forte*, where the clarinet plays a written G, with the instrument all open. This has two implications: it frees the hands of the musician, giving way for wide clarinet gestures, and also lowers the output acoustic power of the instrument, forcing the musician to blow it very strongly to achieve the required intensity level. These factors could highlight this concluding point within the excerpt even further for the clarinetists.

These four regions in the score are always related to relevant harmonic transitions and to melodic phrase endings in the excerpt. The first region is a clear preparation to the phrase ending that takes place in the second region, reinforced by the wide upward leap on the clarinet melody, within a Neapolitan sixth subdominant chord. The last three regions are strongly related to the corresponding melodic phrase endings that take place inside each of them, at relevant harmonic

transitions. The long rests after each target note (in bars 12, 16 and 25) also provide unquestionable breathing points for the clarinetists.

The fact that the four regions in the excerpt where the most consistent movement recurrence and bar duration manipulation patterns occurred for this group of musicians are so highly related to key moments in the music structure is a very strong result. It provides good additional evidence to support the hypothesis that the ancillary physical gestures employed by musicians during performances are closely related to their musical intentions towards the musical realisation. In this case, a clear correlation was observed between the occurrence of the most recurrent and prominent clarinet bell gestures, the main expressive note duration manipulations, and structural aspects of the musical excerpt, such as melodic phrasing and punctuation, or harmonic and dynamic transitions.

Therefore, as in the case of the first experiment, the motion and audio analysis in the second experiment led to highly related conclusions, that are in accordance with the musical structure of Brahms' excerpt discussed in this subsection, indicating a clear musical significance in the ancillary movements of the clarinet performers evaluated in this study.

Chapter 6

Conclusion and Future Work

The main objective of this study is to relate the expressive musical intentions of a group of clarinet players to their physical gestures during musical performances, defining an objective methodology for corporal movement analysis. The method developed for movement representation, segmentation and recurrence analysis led to representative clarinet bell gestures and to notable relations between gestural and musical patterns during several performances. The effective scalar representation of movement in time, through the tangential velocity of the clarinet bell, enabled the use of signal processing techniques to analyse the movement and acoustic data in conjunction. A movement recurrence analysis, based on this representation, revealed that most clarinetists execute highly recurrent movements at specific regions in the excerpts, where they indicate clearer musical intentions by the players. An audio analysis, used to further investigate the expressive content of the performances, indicated that the main note duration manipulations were also executed by the clarinetists at these same musical regions.

Regarding the Mozart data-set, significant differences were observed between the two proposed experimental conditions, according to gestural features. A recurrent sequence of planar gestures, more restrained in the metronome cases, was

found in the last three bars of this excerpt, that were shown to be a key musical moment. The analysis of note inter-onset intervals indicated that the most significant and recurrent note duration manipulations also occurred ate the last three bars of the theme. The metronome had a clear restraining effect on the gestural extension and velocity in this musical region, but did not change the location or duration of the gestures, indicating their local relation to the music structure and its associated expressive content.

Regarding the Brahms data-set, the motion recurrence maps for all musicians indicated that the most prominent and recurrent movements occurred around four regions in the excerpt. Recurrent sequences of clarinet gestures were executed quite frequently in these regions, unlike in the rest of the excerpt. The audio analysis also showed that the bars in the excerpt that led the players to the most significant duration manipulations are all located in these same regions in the music. These regions were shown to be related to relevant harmonic transitions and to melodic phrase endings in the excerpt, indicating that the ancillary gestures of the clarinetists are closely related to their musical intentions. This excerpt is longer than the one used in the first experiment, with more complex harmonic and melodic content, providing more room for expressiveness. This results in the occurrence of a more complex and extensive movement recurrence pattern, centred mainly around these four musical regions.

As previously stated, it can be expected that many aspects of music performance remain mostly unchanged over different executions by an expert player, including the musical intentions and the movement patterns. The strong recurrence and relation to the musical structure found in the physical gestures of clarinetists, specially for the more skilled players, corroborate the hypothesis of a musical significance in their ancillary movements, related to their expressive intentions (as indicated by the audio analysis) and important for the desired musical outcome.

Specifically, a correlation was shown between the recurrence pattern of clarinet bell gestures, the main expressive note duration manipulations, and melodic phrasing, dynamics and harmonic transitions, for a substantial data-set.

Previous studies by Wanderley et al. (2005), Desmet et al. (2012), and Caramiaux et al. (2012) have already pointed to a few local relations between gestures and melody in a performance. However, this is the first time a full-scale musical analysis (harmonic, dynamic, melodic and expressive) of two complex excerpts is objectively coupled with a general gestural pattern consistent over a large group of musicians, using a systematic motion analysis method based on a recurrence measure and validated afterwards through an audio analysis. Also of particular relevance is the local parametric analysis defined for gestures, that revealed the high degree of gestural planarity and provided means to compare between excerpts, musicians and experimental conditions, through the use of basic local gestural features. However, since there are only three standard performances by each musician, it is still not possible at this point to perform a statistically significant comparison of features between the individual players in this study.

The gestures were extracted using a single low-level kinematic feature, the tangential velocity of the clarinet bell. Higher-level features were already used for this purpose, but this choice of method was guided by simplicity and the assumption that the clarinet bell motion, taken relative to a static reference point, is a good general indicator of movement for the player, providing a single trajectory from which local features can be extracted and related to the score. Other tools have also been used to measure motion recurrence in music performances, but the method developed here was chosen because it is very straightforward and based on a simple measure of correlation. Also, the resulting recurrence plots can be easily displayed in alignment with the music structure, several plots can be summed directly to analyse the global recurrence over different groups of performances,

and the temporal fluctuation in the location of the recurrent movements can be assessed through the time offset parameter in the maps.

This study is based on an extensive data-set, including 13 clarinetists, two musical excerpts and 100 performances, but the method can still be applied to even more excerpts, players and classes of musical instruments. This could further expand the scope of use and validity of the method, also through the use of additional gestural and acoustic features, leading to broader conclusions on ancillary movements of music performers. Other acoustic features, such as articulation and legato indexes, energy envelope, spectral centroid, pitch variation and spectral flow can be used to further analyse expressive intentions from the acoustic point of view and relate them to the gestural patterns. Higher-level gestural features can also be used, in order to define an objective categorisation of gestures and relate their temporal and spatial characteristics to musical aspects of the performances and individualities of each musician. Additionally, even if the grouping of movement segments into gestures is guided by the criteria of planarity (checked afterwards through PCA) and circularity, this step of the analysis still relies on visual inspection, giving margin for subjective interpretation. In the future this grouping step could be automatised, for instance by computing the cumulative planarity of successive movement segments, in order to quantitatively guide the whole definition of gestures, making the method more solid.

This method could then be incorporated to musical synthesis, recognition, analysis and teaching systems. For instance, understanding the relation between expressive gestural and acoustic features in a performance could lead to an effective set of expressivity parameters and models in musical synthesis systems and digital musical instruments. The analysis of recurrent gestural patterns could also serve as a pedagogical tool for music teachers. Through the comparison with referential performances of experts and previous performances by the students it is

66

possible to evaluate many expressive aspects of their executions and their evolution over time. The tools and results presented in this study can also be used in the development of new theories in the fields of musicology, human cognition and physiology, ultimately defining a musical meaning for the physical gestures of musicians, and answering important questions, such as: Is it possible to effectively correlate movement and sound?

Bibliography

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

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*, 131(3):2162–2172, 2012.

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*, pages 24–27, 2009.

A. Camurri, G. De Poli, M. Leman, and G. Volpe. A multi-layered conceptual framework for expressive gesture applications. *Proceedings Intl EU-TMR MOSART Workshop*, 2001.

A. Camurri, B. Mazzarino, M. Ricchetti, R. Timmers, and G. Volpe. Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri and G. Volpe, editors, *Gesture-Based Communication in Human-Computer Interaction*, Lecture Notes in Computer Science, pages 20–39. Springer, Berlin, 2004.

B. Caramiaux, M. Wanderley, and F. Bevilacqua. Segmenting and parsing instrumentalists' gestures. *Journal of New Music Research*, 41(1):13–29, 2012.

S. Dahl and A. Friberg. Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24(5):433–454, 2007.

J. W. Davidson. Visual perception and performance manner in the movements of solo musicians. *Psychology of Music*, 21:103–113, 1993.

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

J.W. Davidson. What does the visual information contained in music performances offer the observer? some preliminary thoughts. In R. Steinberg, editor, *Music and the mind machine: Psychophysiology and psychopathology of the sense of music*, pages 105–114. Heidelberg: Springer., 1995.

G. De Poli, S. Canazza, C. Drioli, A. Rodà, and A. Vidolin. Modeling and control of expressiveness in music performance. *Proceedings of the IEEE*, 92(4):686–701, 2004.

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*, 41(1):31–48, 2012.

D. Fenza, L. Mion, S. Canazza, and A. Rodà. Physical movement and musical gestures: A multilevel mapping stategy. *Proceedings of Sound and Music Computing*, pages 1–6, 2005.

A. Gabrielsson. Expressive intention and performance. In Steiner, editor, *Music and the Mind Machine*, pages 35–47. Springer, Berlin, 1995.

A. Gabrielsson. Music performance research at the millenium. *Psychology of Music*, 31:221–272, 2003.

P. Juslin. Emotional communication in music performance: A functionalist perspective and some data. *Music Perception*, 14:383–418, 1997.

P. Juslin. Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human perception and performance*, 26:1797–1813, 2000.

M. Leman. *Embodied music cognition and mediation technology.* The MIT Press, Cambridge, MA, 2007.

E. Maestre and E. Gómez. Automatic characterization of dynamics and articulation of expressive monophonic recordings. *Audio Engineering Society 118th Convention Paper*, page 8, 2005.

D. McNeill. *Gesture and Thought.* University Of Chicago Press, Chicago, 2007.

L. Naveda and M. Leman. A cross-modal heuristic for periodic pattern analysis of samba music and dance. *Journal of New Music Research*, 38(3):255–283, 2009.

N. Rasamimanana. Towards a conceptual framework for exploring and modelling expressive musical gestures. *Journal of New Music Research*, 41(1):3–12, 2012.

P. Senin. Dynamic time warping algorithm review. Technical report, Information and Computer Science Department, University of Hawaii, Honolulu, USA, 2008.

E. Teixeira, M. Loureiro, and H. Yehia. Methodological aspects of the research in musical expressiveness based on corporal movement information. Unpublished Report. Retrieved from http://hal.archives-ouvertes.fr/hal-00611660/en/, 2010.

E. Teixeira, M. Loureiro, H. Yehia, and M. Wanderley. Motion recurrence analysis in music performances. *Proceedings of the Sound and Music Computing Conference*, pages 317–322, 2013.

E. Teixeira, M. Loureiro, M. Wanderley, and H. Yehia. Motion analysis of clarinet performers. *Journal of New Music Research*, 43:1–15, 2014.

G. Varni, G. Volpe, and A. Camurri. A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. *IEEE Transactions on multimedia*, 12(6):576–590, 2010.

B. Vines, I. Dalca, and M. Wanderley. Variation in expressive physical gestures of clarinetists. *Proceedings of the 9th International Conference on Music Perception and Cognition*, pages 1721–1722, 2006a.

B. Vines, C. Krumhansl, M. Wanderley, and D. Levitin. Cross-modal interactions in the perception of musical performance. *Cognition*, 101:80–113, 2006b.

M. Wanderley. Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth and T. Sowa, editors, *Gesture and Sign Language in Human-Computer Interaction*, Lecture Notes in Computer Science, pages 241–253. Springer, Berlin, 2002.

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*, 34(1):97–113, 2005.

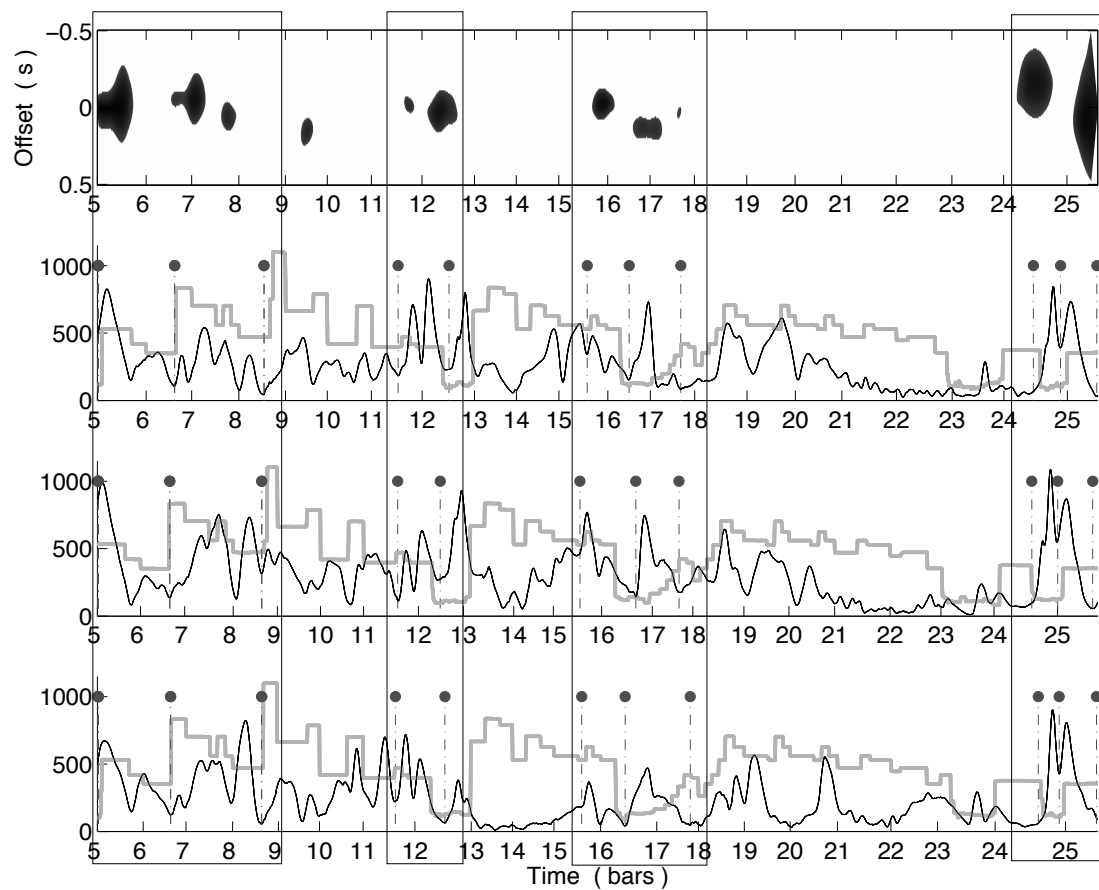# Appendix A

Examples of Recurrent Gestures
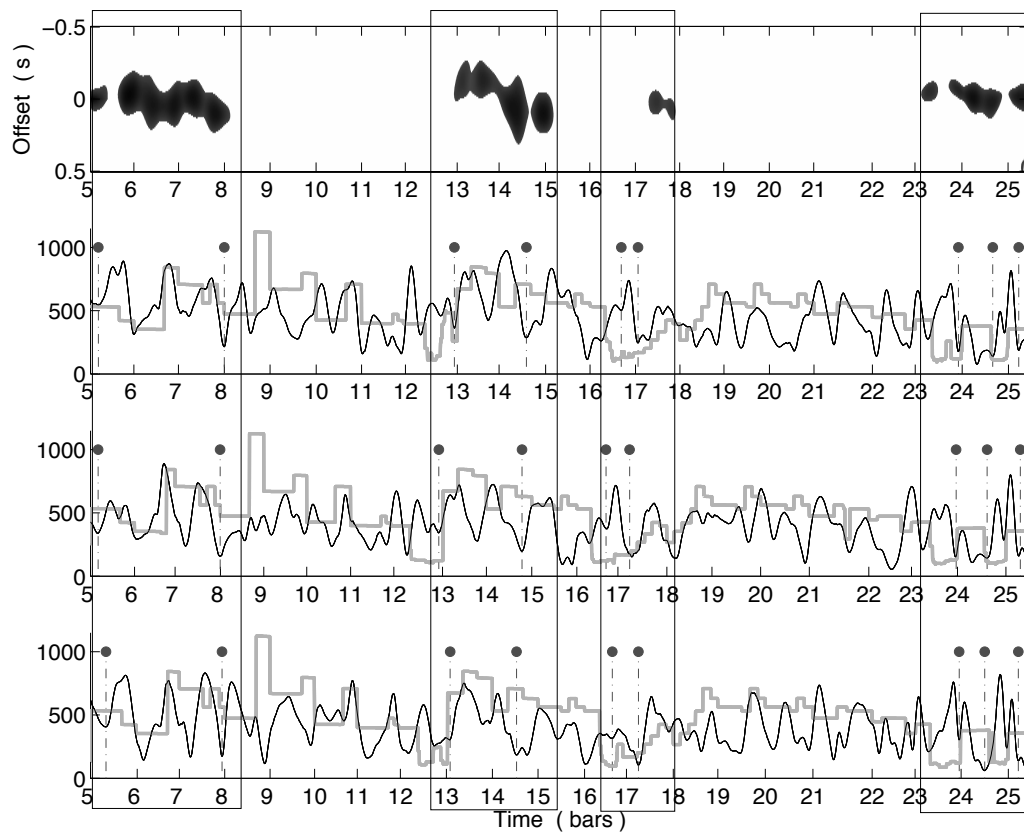
Figure A.1: Regions of interest for the three performances of Brahms' excerpt by Musician 3, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map.
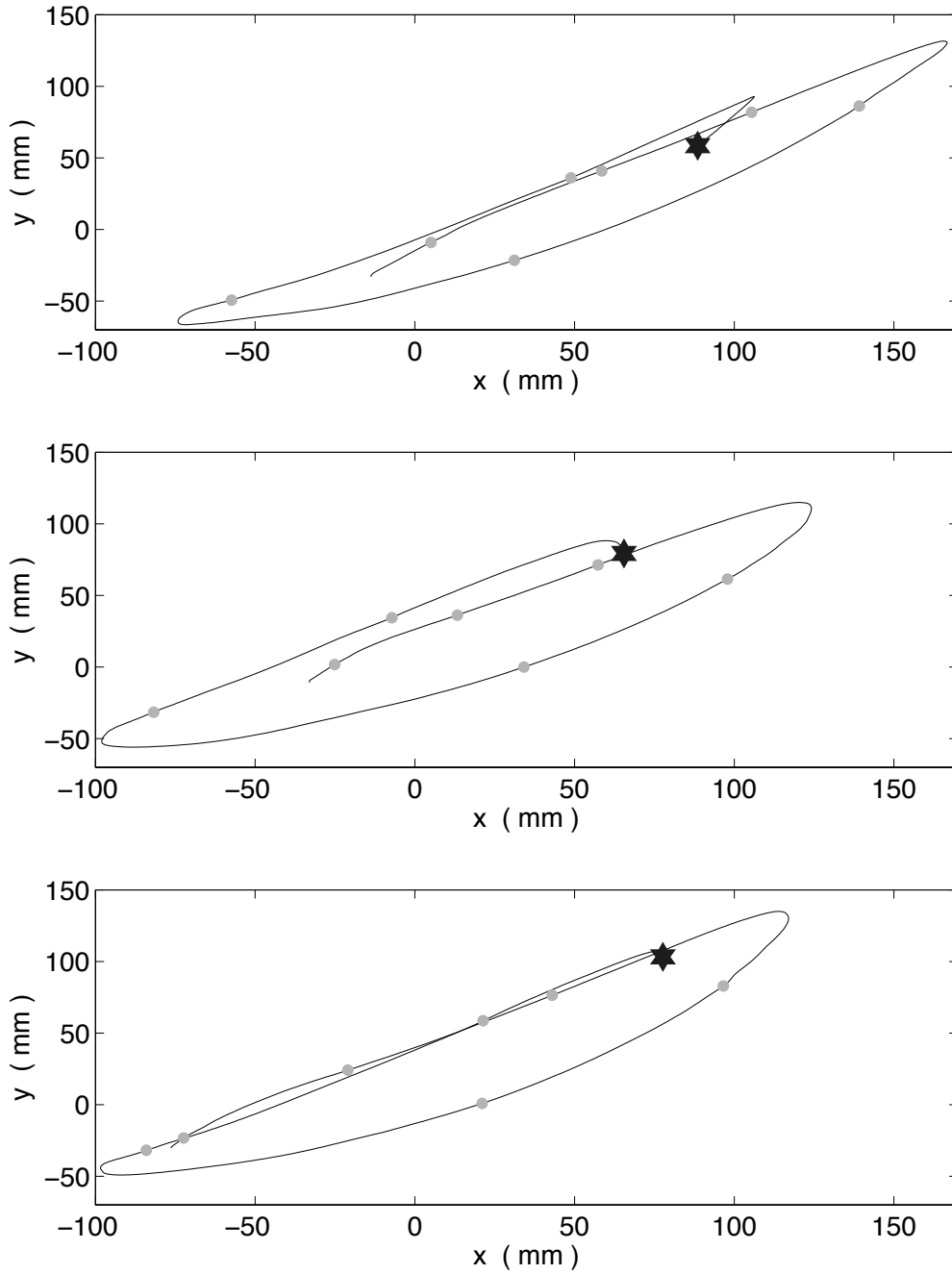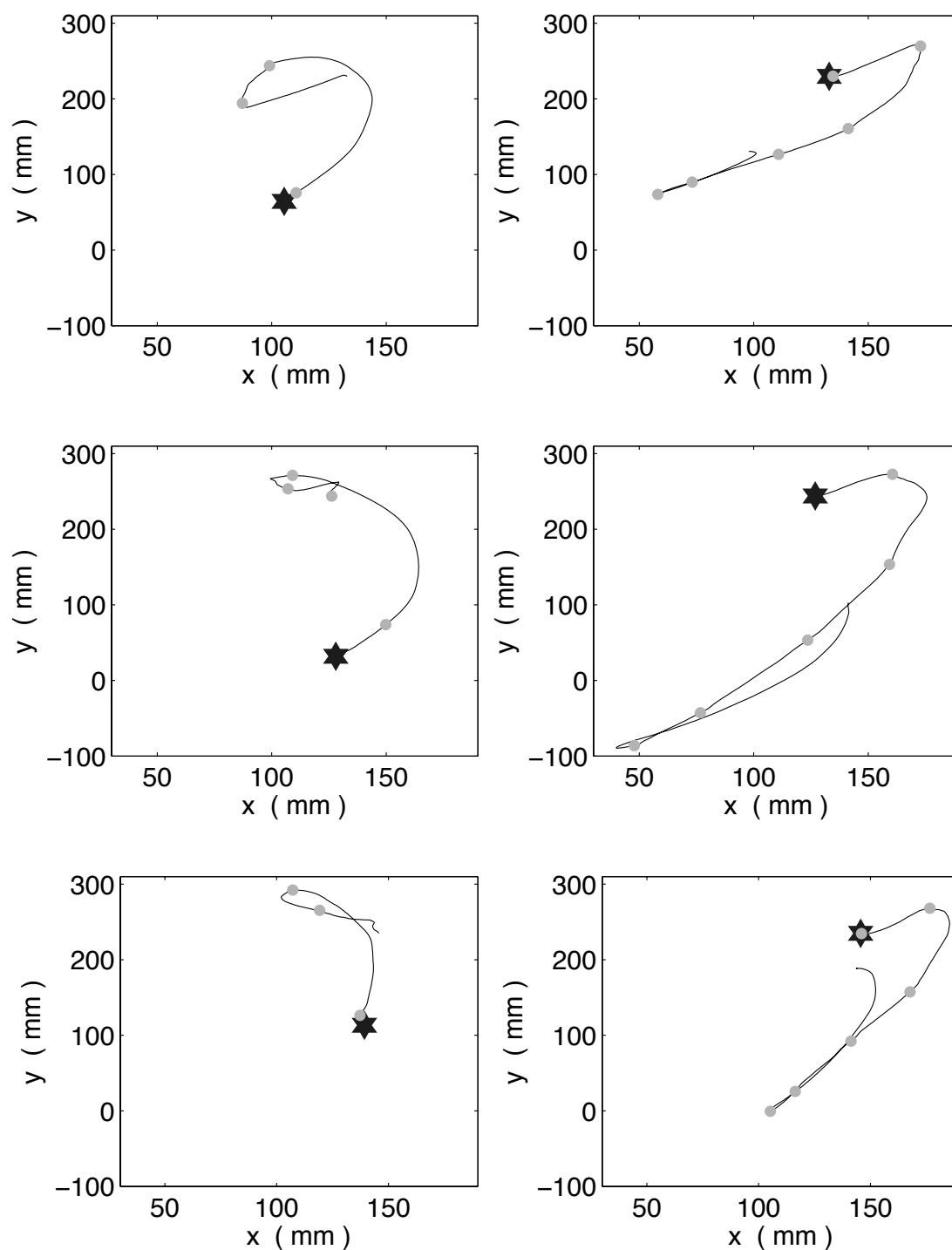
Figure A.2: 2D trajectories of the gesture inside the first region of interest for the three performances of Brahms' excerpt by Musician 3. The star marks the initial point and the circles indicate the note onsets.
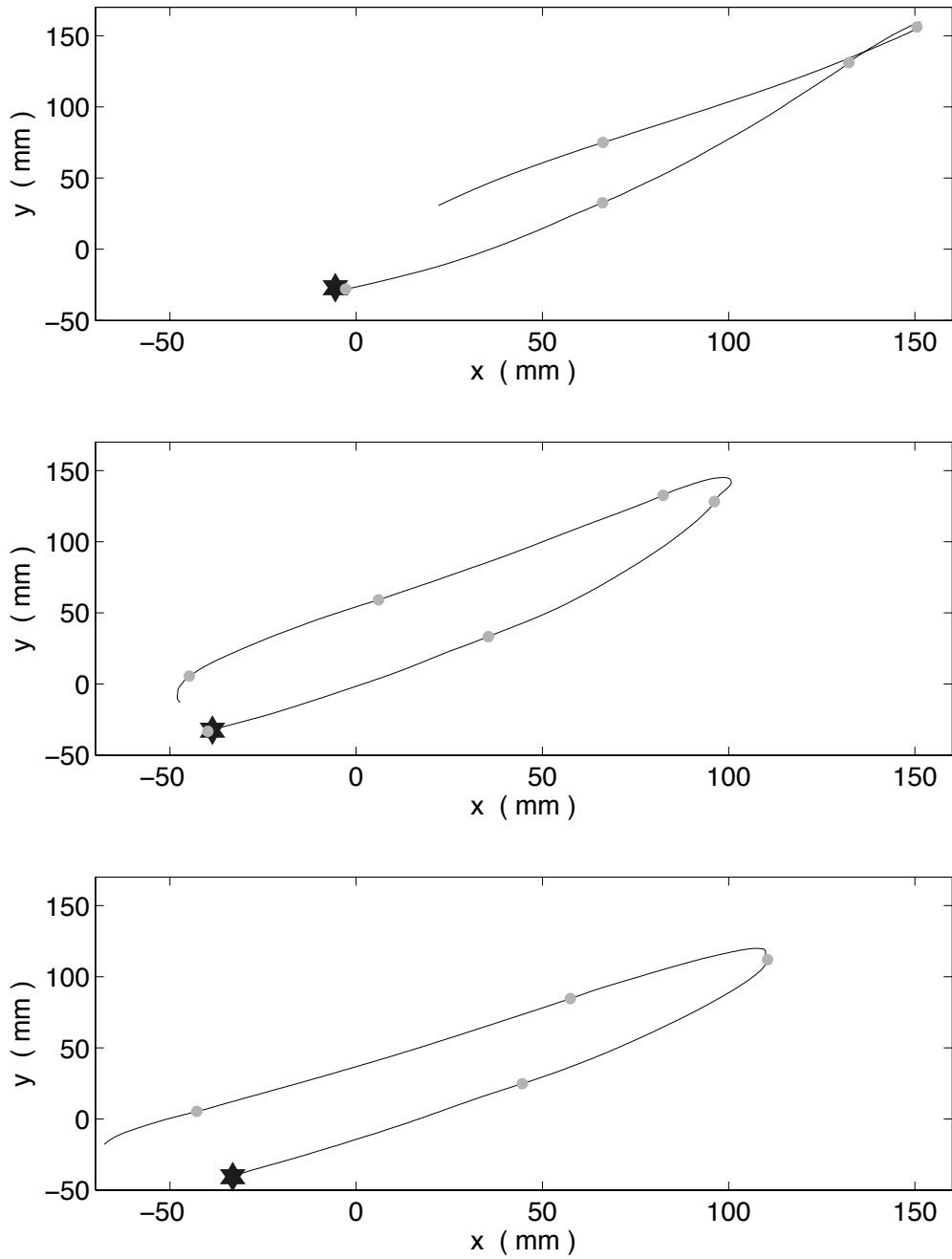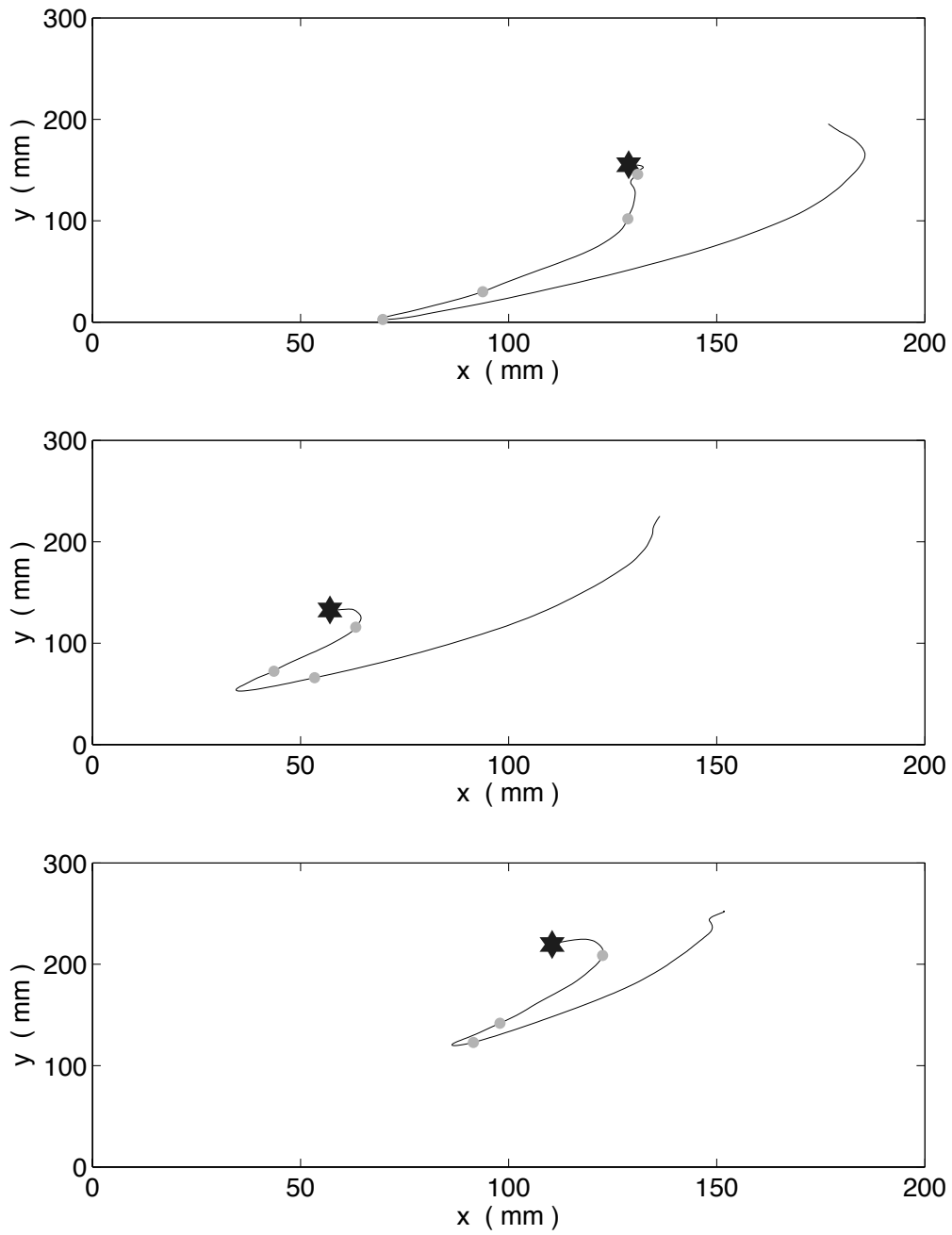
Figure A.3: 2D trajectories of the gesture inside the second region of interest for the three performances of Brahms' excerpt by Musician 3. The star marks the initial point and the circles indicate the note onsets.
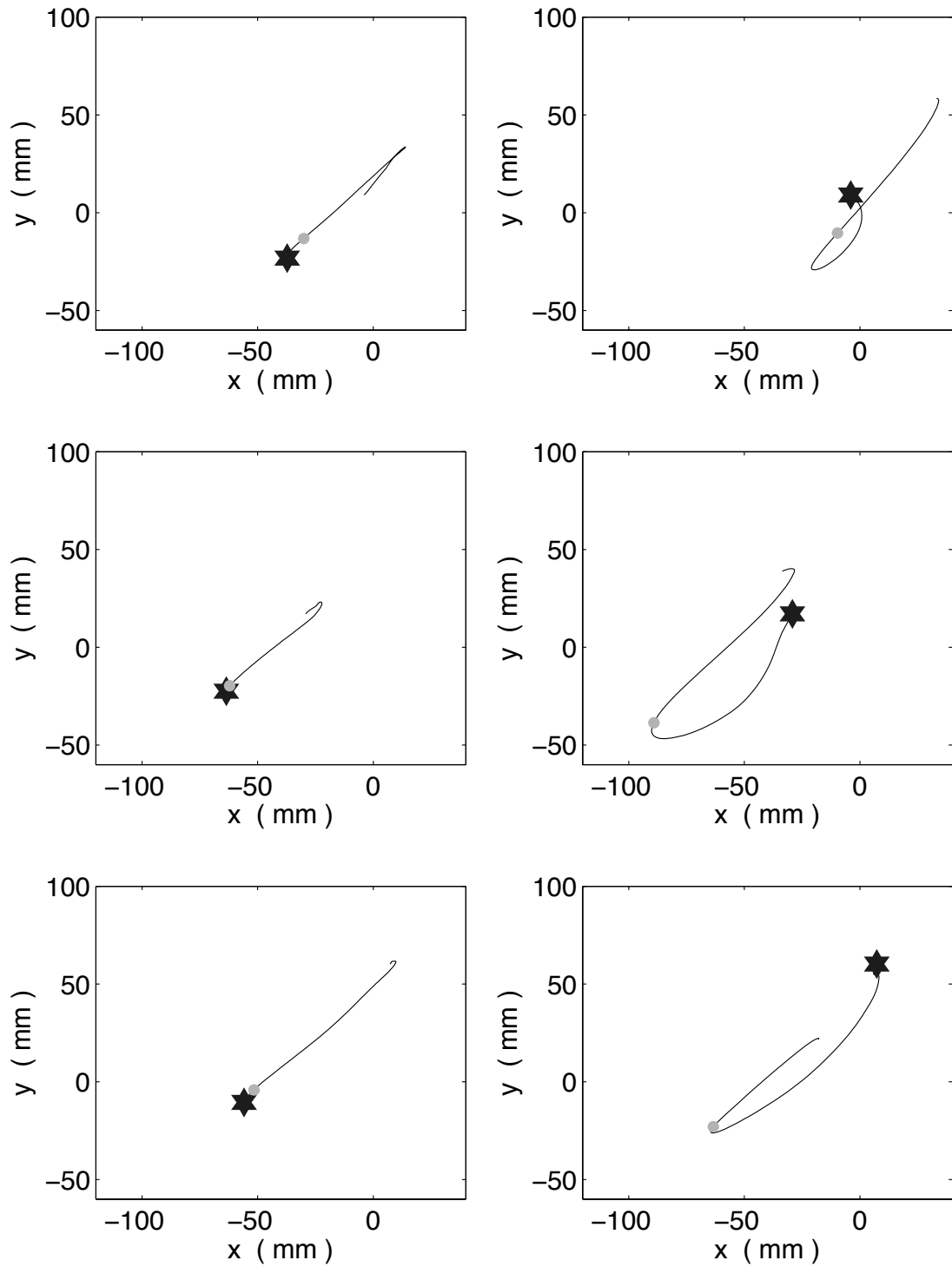
Figure A.4: 2D trajectories of the two gestures inside the last region of interest for the three performances of Brahms' excerpt by Musician 3. The star marks the initial point and the circles indicate the note onsets.
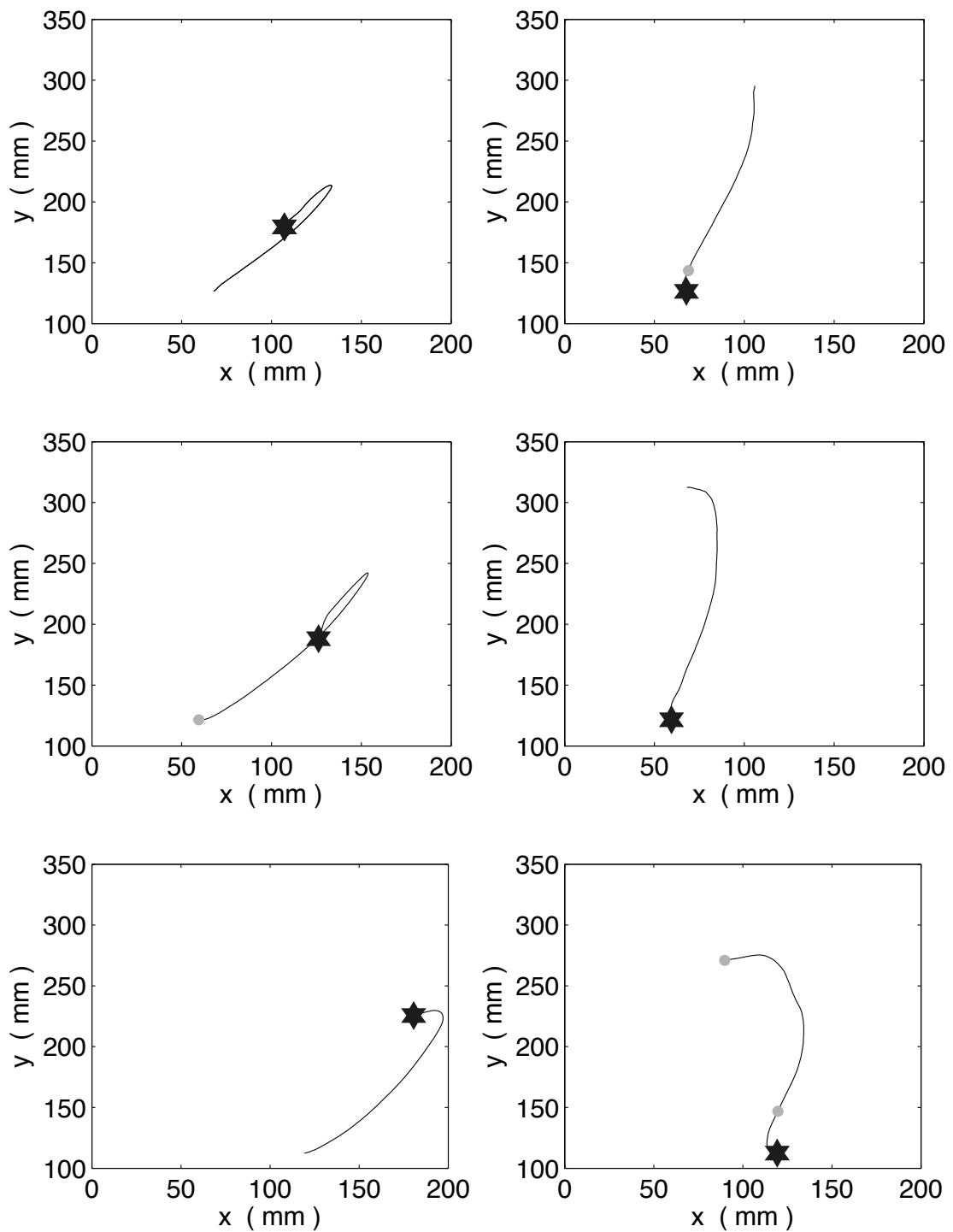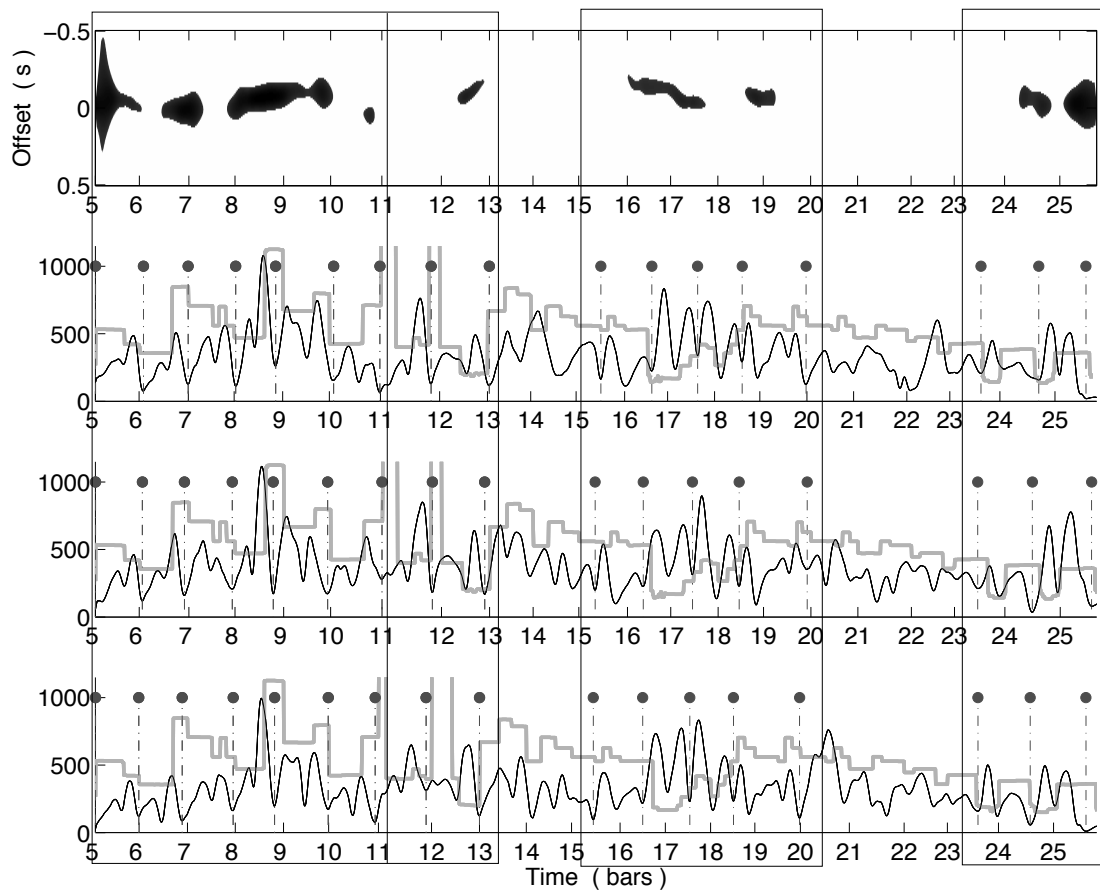
Figure A.5: Regions of interest for the three performances of Brahms' excerpt by Musician 8, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map.
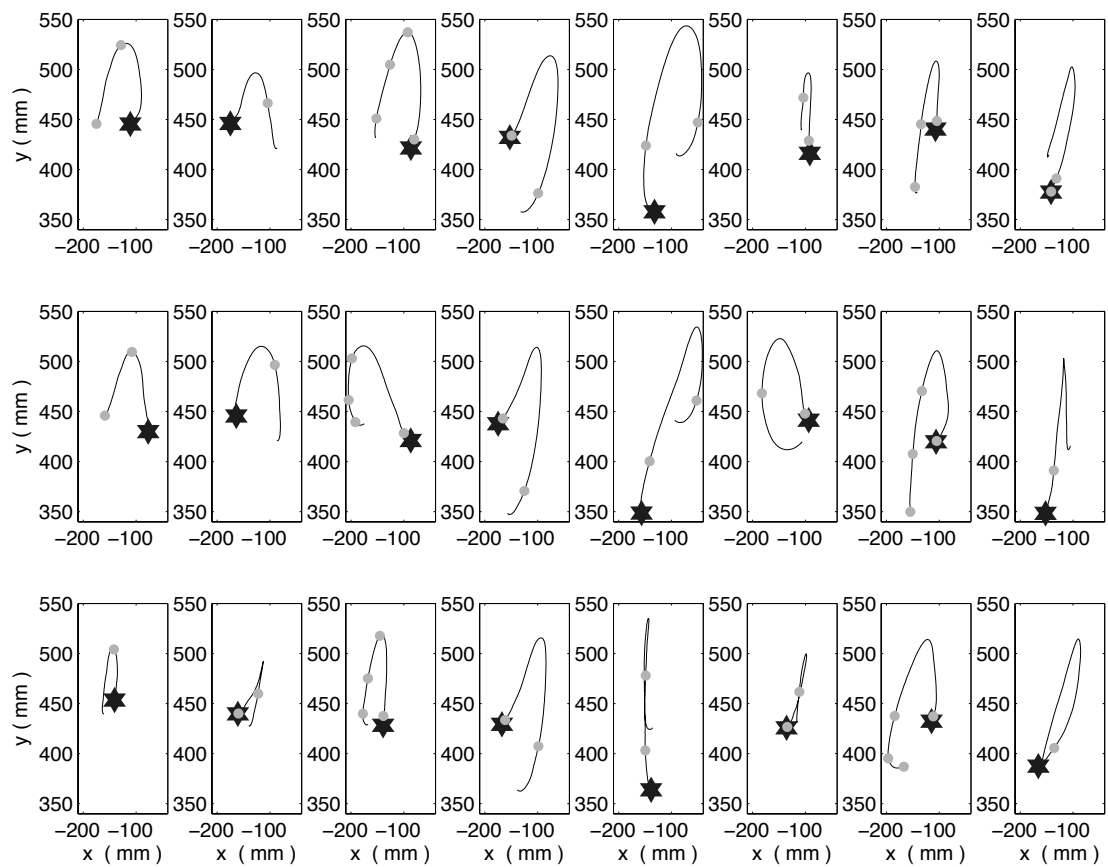
Figure A.6: 2D trajectories of the two gestures inside the first region of interest for the three performances of Brahms' excerpt by Musician 8. The star marks the initial point and the circles indicate the note onsets.

Figure A.7: 2D trajectories of the gesture inside the second region of interest for the three performances of Brahms' excerpt by Musician 8. The star marks the initial point and the circles indicate the note onsets.

Figure A.8: 2D trajectories of the two gestures inside the last region of interest for the three performances of Brahms' excerpt by Musician 8. The star marks the initial point and the circles indicate the note onsets.
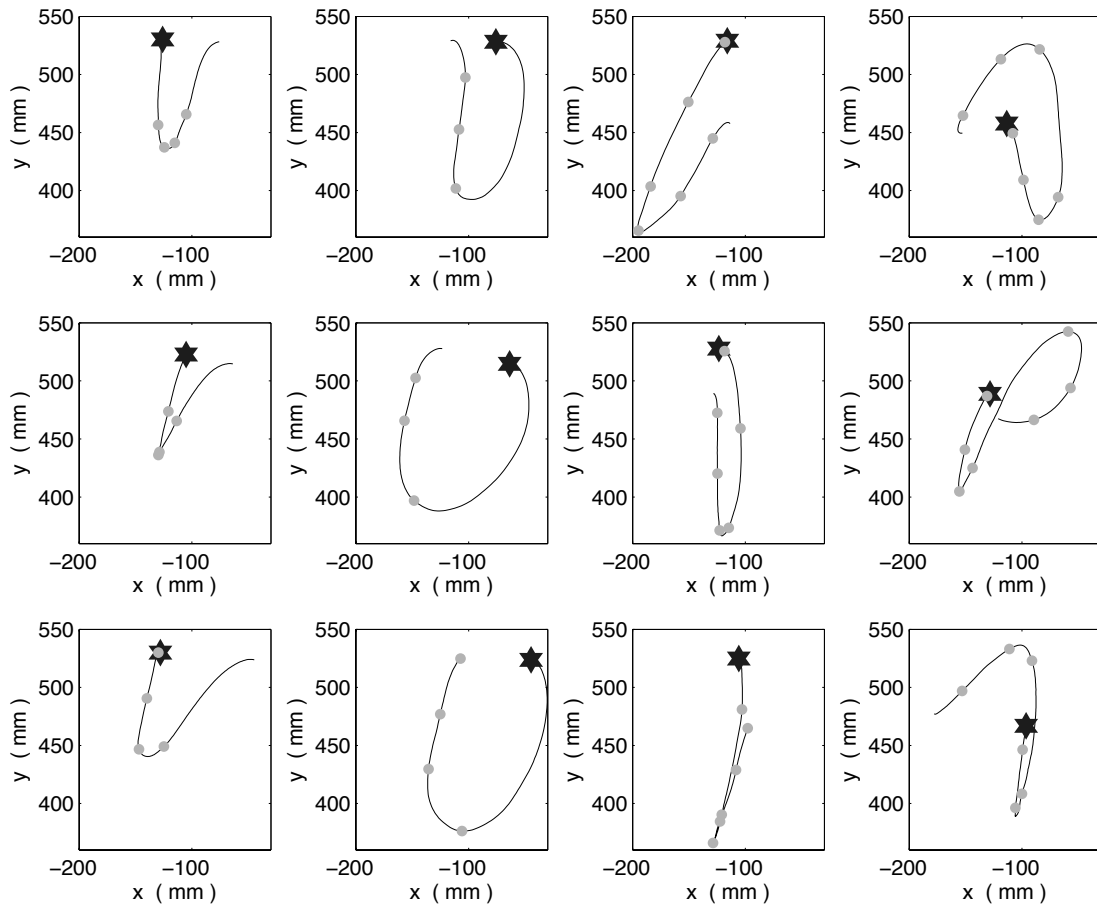
Figure A.9: Regions of interest for the three performances of Brahms' excerpt by Musician 12, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map.

Figure A.10: 2D trajectories of the eight gestures inside the first two regions of interest for the three performances of Brahms' excerpt by Musician 12. The star marks the initial point and the circles indicate the note onsets.
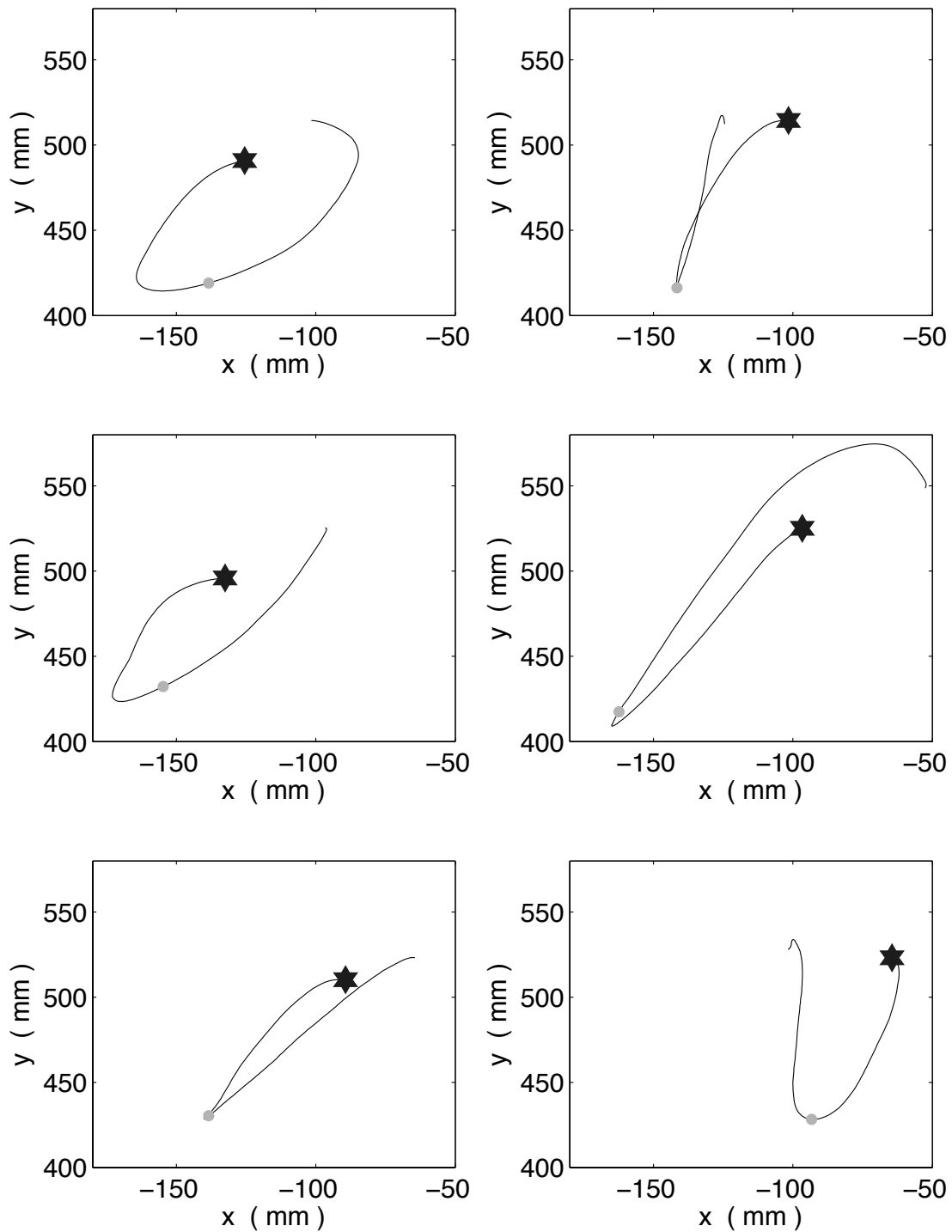
Figure A.11: 2D trajectories of the four gestures inside the third region of interest for the three performances of Brahms' excerpt by Musician 12. The star marks the initial point and the circles indicate the note onsets.

Figure A.12: 2D trajectories of the two gestures inside the last region of interest for the three performances of Brahms' excerpt by Musician 12. The star marks the initial point and the circles indicate the note onsets.
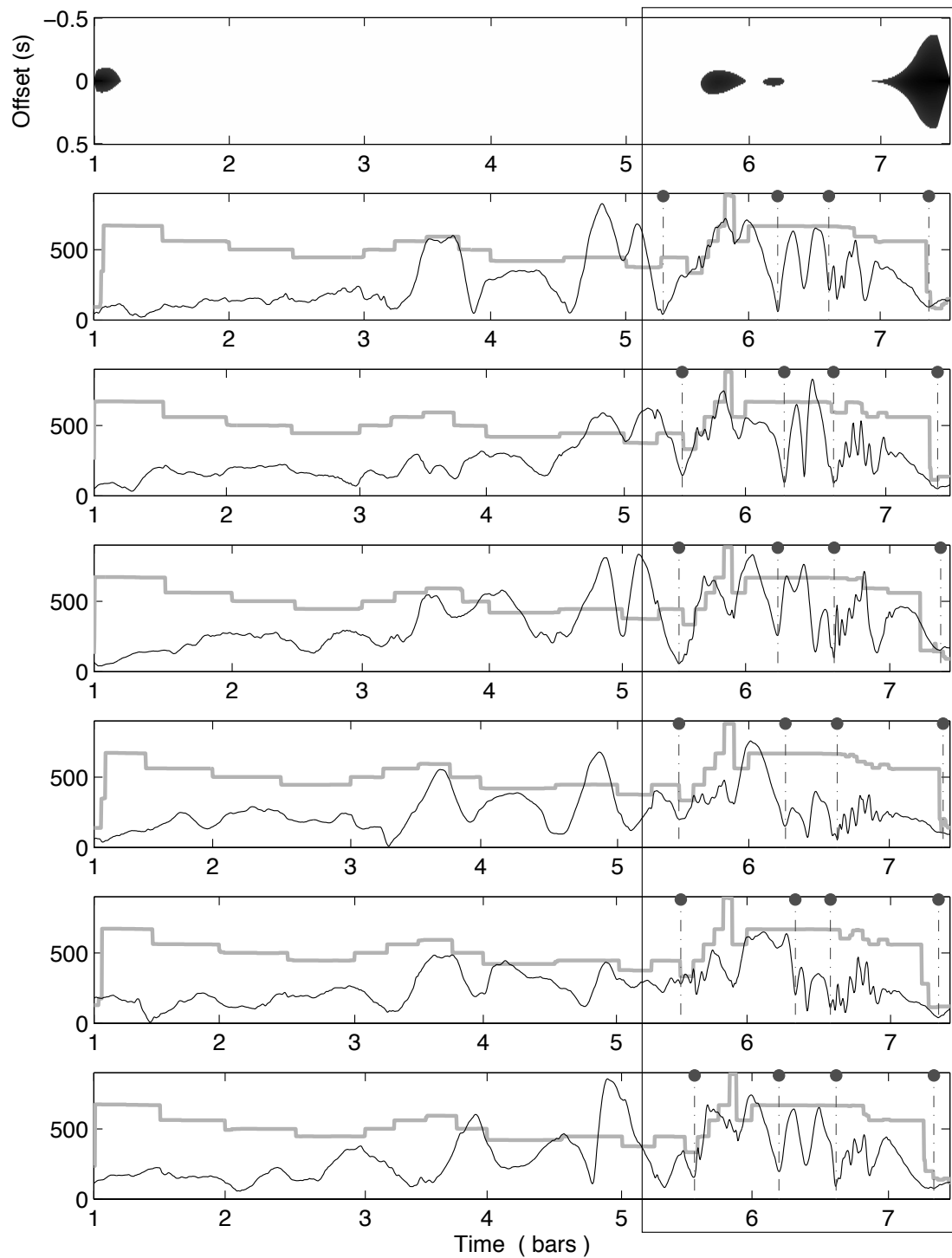
Figure A.13: Regions of interest for the six performances of Mozart's excerpt by Musician 4, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map. Metronome performances at the three bottom plots.
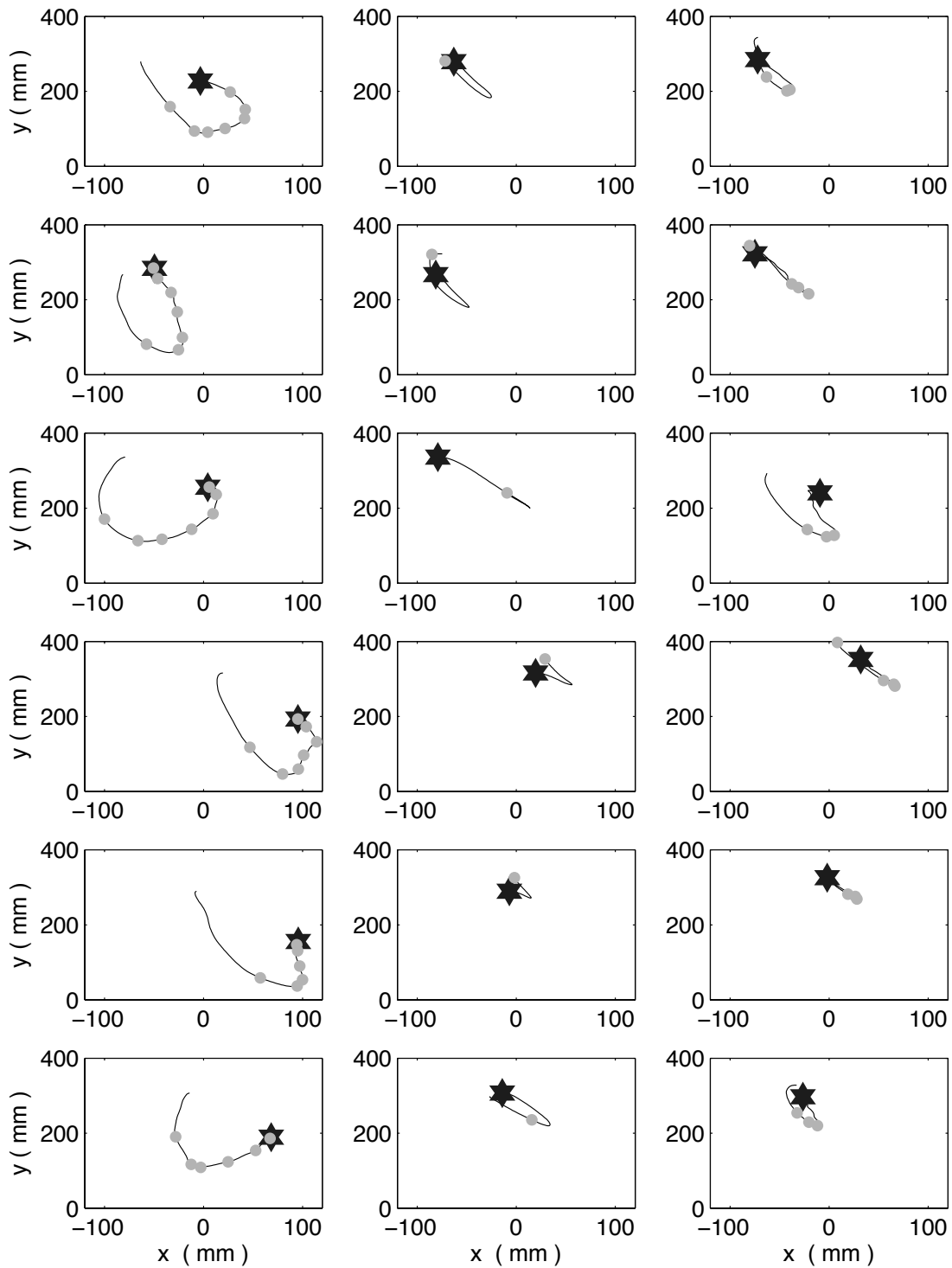
Figure A.14: 2D trajectories of the three gestures inside the region of interest for the six performances of Mozart's excerpt by Musician 4. The star marks the initial point and the circles indicate the note onsets. Metronome performances at the three bottom rows.
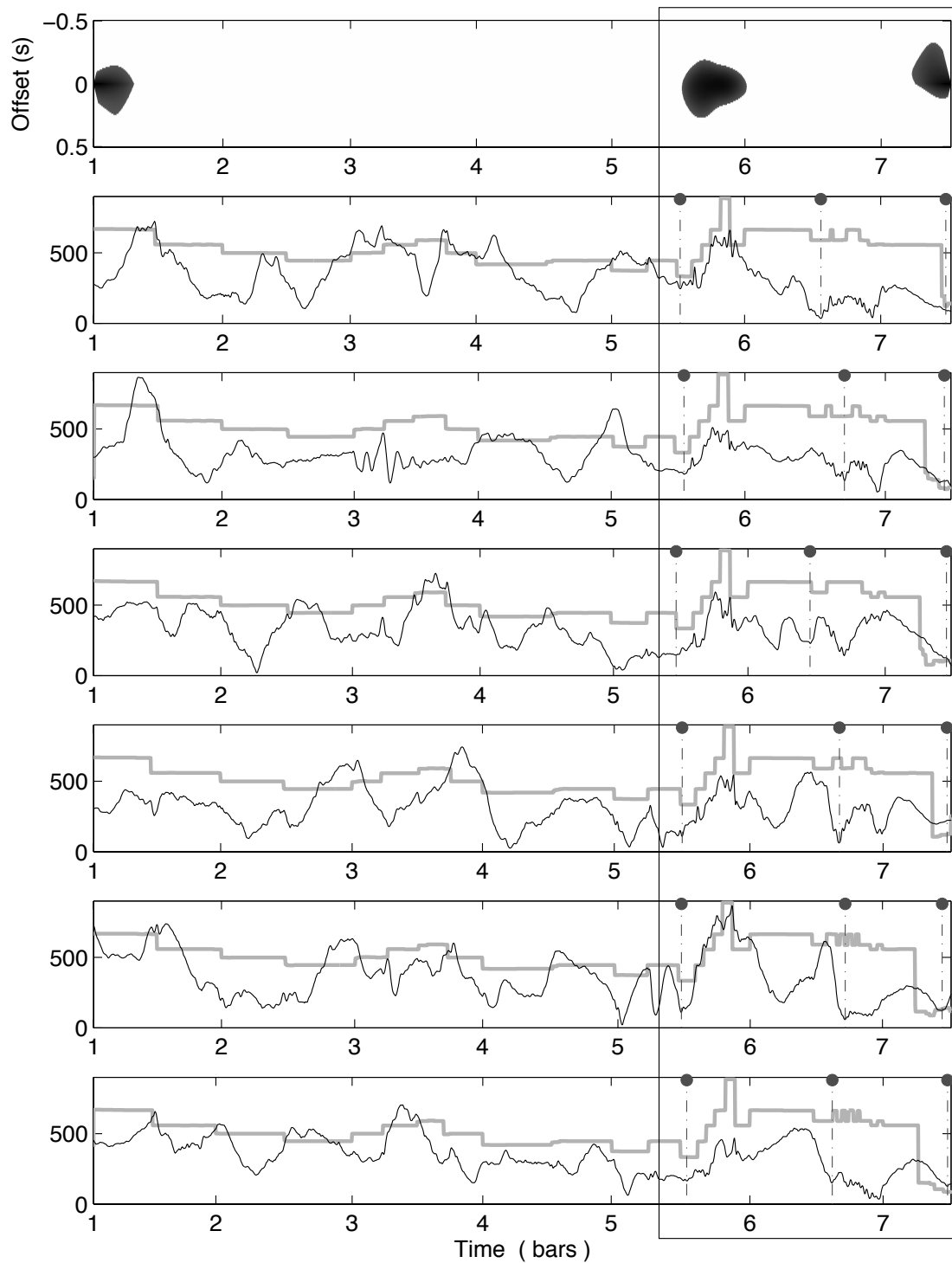
Figure A.15: Regions of interest for the six performances of Mozart's excerpt by Musician 7, and their constituent gestures, marked by the dashed vertical lines. The grey curves show the pitch (Hz), the black curves show the tangential velocity (mm/s), and the top plot shows the motion recurrence map. Metronome performances at the three bottom plots.
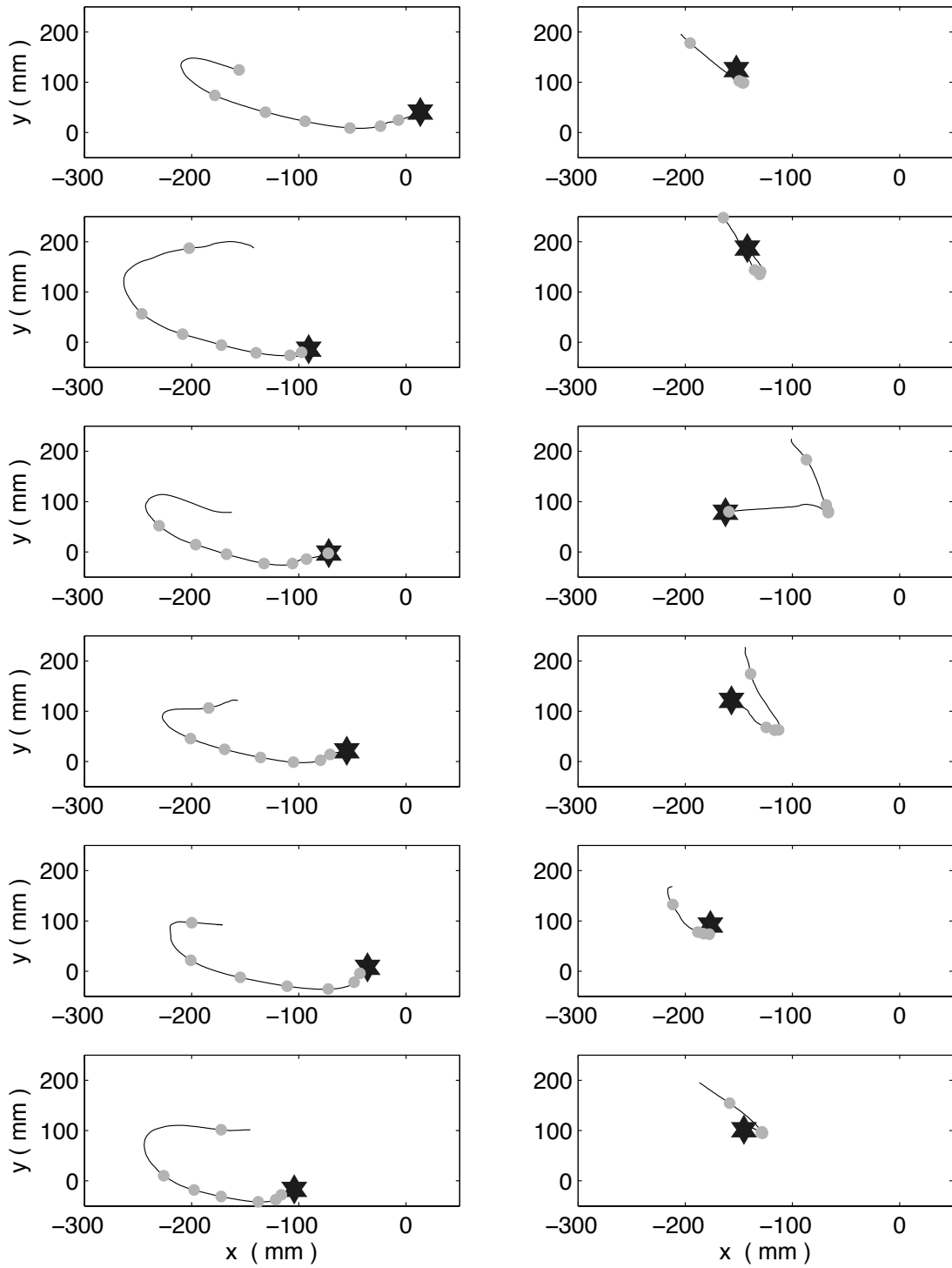
Figure A.16: 2D trajectories of the two gestures inside the region of interest for the six performances of Mozart's excerpt by Musician 7. The star marks the initial point and the circles indicate the note onsets. Metronome performances at the three bottom rows.