Faculté Polytechnique

Development of an optical motion capture setup for feature extraction and statistical analysis of the pianist’s expert gestures

Master’s Thesis
Completed at CIRMMT, McGill University, Montreal (QC), Canada

Submitted in partial fulfillment of the requirements for the degree of Master of Electrical Engineering, specialized in Multimedia and Telecommunications

Mickaël Tits

Under the supervision of: Prof. Marcelo WANDERLEY (McGill University, Montreal (QC), Canada), Prof. Thierry DUTOIT, Dr. Nicolas D’ALESSANDRO, and Dr. Joëlle TILMANNE (University of Mons, Belgium)

Academic year 2013-2014
Abstract

Movements of the body and especially of the hand are extremely complex, and are driven by several processes of biomechanical and neuroplastic types, among others. Their study presents a great interest in Human-Computer Interaction (HCI), in sportive and musical education, as well as in medicine and reeducation. The aim of this research was to capture and analyze highly complex gestures performed by experts in hand motor control, such as musicians. Musical gestures are highly formalized by indications generally noted on a musical score, allowing a more focused and thus objective study of the hands movements. Piano playing, which requires rather precise and coordinated movements, was therefore chosen as our case study.

In this research, we developed an optical motion capture setup that allowed the capture of pianists’ expert gestures in playing complex piano pieces. The setup consisted of 12 infrared cameras aimed at the keyboard and capturing in 3D all Degrees Of Freedom (DOF) of both hands, by means of 26 retro-reflective markers placed on each hand. With this setup, we captured performances by four pianists with different expertise levels, playing six different piano pieces and an exercise. An example of the successful capture of the Fantasy in C Major by Robert Schumann can be found here: http://youtu.be/lNwWHlkjMJ4.

The exercise allowed us to analyze redundant DOF solicitations in a basic keystroke gesture according to different tempi and movement instructions. This analysis showed that pianists have their own movement signature in natural playing. It also showed that the same pianist can use its redundant DOF of the arm and hand differently, and thus play the same musical pattern in countless different ways. We also carried out a Principal Component Analysis (PCA) of the hands gestures, based on the six pieces performed by the four pianists. Results of PCA showed the power of this tool for dimensionality reduction, and representation of pianists’ expert gestures. Eight Principal Components (PC) were generally sufficient to represent 95% of the 23 DOF of the hand (and only three PC for 80%). Furthermore, PCA allowed us to decompose hands movements into eigengestures associated with each PC (see http://youtu.be/6qIh99UuJ4). Finally, both analyses of the exercise and of the piano pieces seemed to show that a more trained pianist has a finer hand motor control, and in particular, a finer balance between the two hands.

Results of this study show that PCA is a powerful tool for motion analysis and could be applied in different fields such as HCI, and sportive and musical education. Furthermore, in a clinical context, PCA could provide an assessment tool for patients suffering from motor disability.

Keywords: optical motion capture, musical gesture, hand biomechanics, Principal Component Analysis (PCA)

Contact: Mickaël Tits, University of Mons - mickaeltits@gmail.com
Foreword

This master’s thesis has been conducted at the Center for Interdisciplinary Research in Music Media and Technology (CIRMMT) of McGill University (Montreal, Canada). This project was multidisciplinary, and involved disciplines such as music theory, hand biomechanics, motion capture and signal processing. Its implementation could be done thanks to a collaboration with the CIRMMT and the Numediart Institute from the University of Mons (Belgium). My own background in music theory and piano added a grain of salt in the numerous different disciplines required for the completion of this thesis.

This project was supervised by Professor Marcelo Wanderley, director of the CIRMMT. I would like to warmly thank him, for his support and patience, and for the numerous contacts he gave me for the achievement of this project. I also would like to thank Doctor Nicolas d’Alessandro and Doctor Joëlle Tilmanne, for their support from Mons, and the guidance they gave me throughout the entire project. Furthermore, I would like to thank Professor Thierry Dutoit for providing me the opportunity to achieve this valuable international experience. Finally, I thank my family for supporting me during the entire project, and for their help in the proofreading of my report.

The idea of this project was born from discussion with my supervisors about interaction devices, signal processing, and piano. Research on input devices in the context of interaction with machines is recently very active, and often uses movements of the hand to achieve these interactions. The hand is a very sophisticated mechanism, and its high dimensionality makes it difficult to capture, analyze and model. In the present project, I developed a motion capture setup to fully capture movements of the hands performing highly complex gestures, that are pianists’ expert gestures. Then, thanks to that developed setup, I analyzed the hand motor control of different pianists. Results of my analysis may benefit to applications in human-computer interaction and in biomechanics and medicine.
# List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMI</td>
<td>Digital Musical Instrument</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>CIRMMT</td>
<td>Center for Interdisciplinary Research in Music Media and Technology</td>
</tr>
<tr>
<td>MIDI</td>
<td>Musical Instrument Digital Interface</td>
</tr>
<tr>
<td>DOF</td>
<td>Degree Of Freedom</td>
</tr>
<tr>
<td>ROM</td>
<td>Range Of Movement</td>
</tr>
<tr>
<td>CMC</td>
<td>Carpometacarpal</td>
</tr>
<tr>
<td>MCP</td>
<td>Metacarpophalangeal</td>
</tr>
<tr>
<td>DIP</td>
<td>Distal Interphalangeal</td>
</tr>
<tr>
<td>PIP</td>
<td>Proximal Interphalangeal</td>
</tr>
<tr>
<td>FPS</td>
<td>Frames Per Second</td>
</tr>
<tr>
<td>BPM</td>
<td>Beats Per Minute</td>
</tr>
<tr>
<td>MOCAP</td>
<td>Motion Capture</td>
</tr>
<tr>
<td>IOI</td>
<td>Inter-Onset Intervals</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PC</td>
<td>Principal Component</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>ASD</td>
<td>Angle Standard Deviation</td>
</tr>
<tr>
<td>CER</td>
<td>Cumulative Eigenvalue Ratio</td>
</tr>
</tbody>
</table>
Introduction

Motivations

Music has always been a discipline of great interest in research, education and therapy, because of its virtues on the motor control of the body, and on the arousal of auditory and emotional brain activities. For centuries, music has always been closely linked with the advances in science and technology. The piano itself is a marvel of mechanical technology, allowing production of highly complex music, created by the coordinated movements of both hands and feet. The improvement of acoustical instruments is tightly connected with research in mechanical and acoustical physics. More recently, the development of electronic music has followed the progress in computer science.

A recent research field in music is the design of new Digital Musical Instruments (DMI) [1], making use of Human-Computer Interaction (HCI) technologies to enhance the interactions between the musician and the produced sound. The object of this field of research is to extend the creative potential of existing instruments by adding new parameters, captured with additional sensors, into the equation governing the causal connection between the performers' gestures and the created music. In other words, the design of new DMI aims for the creation of complete and optimized instruments, by amplifying the impact of the users’ gestures on the music. The control subtlety allowed by the introduction of new parameters to the instruments brings new possibilities in music composition with increased performances.

Research in the design of DMI is highly multidisciplinary. First of all, it involves the study of music theory and sound synthesis. But another important part is the gestural control of the instrument. The quality of the gestural control of an instrument first depends on the mapping with the sound synthesis, using motion capture and gesture modeling and recognition. Secondly, the different movements used for the control of the instrument have a great impact on the complexity of the latter, and on the possibilities of expression of the performer. Before the design of an instrument, it is thus important to have a specific understanding of the operation of the performer’s gestures. Another important discipline is therefore involved in the design of DMI: biomechanics, i.e. the study of the anatomy and the motor control of the limbs used for interaction with the instrument.

The hand is the most used body limb for interactions with objects, because of its exceptional mechanics, allowing the execution of millions of different complex gestures. Its operation depends on muscles, articulations, and neurons, whose actions allow for a formidable range of movements. These actions are optimized for the simple execution of the most important gestures in everyday life, like prehension, and can therefore limit the extend of the gestures of the hand.

This impressive and complex "biological technology" that the hand is must thus be correctly understood to allow an adequate association with the control of machines. A reliable way to do in-depth study of the operation of the hand is to study complex gestures performed by experts. For the particular case of the hand, and in the context of interactions with machines, a particularly suited discipline for the study of expert gestures is music. Musicians must indeed be experts in the motor control of the hands, as the playing of music requires highly complex, precise and coordinated gestures involving all articulations of the hands. Additionally, the discipline of music also has a major advantage for the study of hand gestures, as it is a highly formalized discipline. Indeed, a musician playing an instrument
always follows very precise indications and techniques, generally noted on a musical score. This feature of the music discipline allows thus a focus of the study on objective gestures.

Expert gestures in piano playing are especially interesting to study, as they require tremendous coordination of both hands and all fingers, and involve a number of varied and highly complex gestures. Many studies already investigated pianists’ expert gestures, but most of them were limited to the capture of a simple piano exercise on a single hand, and calculation of simple statistical values. Their limitations were often due to the fact that the study of musicians’ expert gestures requires the combination of several skills and disciplines that are music theory, biomechanics, motion capture, and signal processing.

The motivation for the achievement of this project was to conciliate all the skills and disciplines that are necessary for an effective study of the subject. This was made possible by a collaboration with the Center for Interdisciplinary Research in Music Media and Technology (CIRMMT) of McGill University, and the Numediart Institute of the University of Mons. This multidisciplinary project led to two main innovations in the analysis of pianists’ expert gestures. First, a motion capture setup has been developed to allow the complete and precise capture of both hands on a complex piano piece. The second innovation, achievable thanks to the data captured with the aforementioned motion capture setup, was a motion analysis of actual piano pieces, rather than limited exercises. This allowed an in-depth study of real complex and varied gestures. This analysis was conducted via the use of the powerful tool that is the Principal Component Analysis (PCA). Additionally to these innovations, a "per-finger" piano exercise has been developed, allowing analysis of the control of each finger separately. Finally, a first well-supplied database of both hands trajectories for several pianists on different pieces has been established. Moreover, this database has been furnished with synchronized MIDI data that can serve as landmarks and allow annotations.

Project Description

This report is divided in three parts.

- Part I aims to give the reader the necessary background for a proper understanding of the conducted research. This part is composed of three chapters. In Chapter 1, we first present a review of related work, concerning motion capture and motion analysis. Chapter 2 briefly outlines the theory on the hand biomechanics that is necessary to fully apprehend the study completed in the present research. The tools used in this research are then presented in Chapter 3.

- Part II is composed of two chapters and presents the motion capture part of the present research. In Chapter 4, we explain the development of a setup allowing the capture of pianists’ expert gestures on an entire piano piece. Chapter 5 then describes the motion capture database collected in the framework of the present thesis. This database contains motion capture recordings of a piano exercise and several pieces.

- Part III is composed of two chapters focusing on the analysis of pianists’ expert gestures. In Chapter 6, we study pianists’ motor control of the hand by analyzing data captured in the course of an exercise specifically developed for this purpose. Finally, Chapter 7 presents a PCA of pianists’ expert gestures in actual performance.

Final conclusions are then presented at the end of the report.
Part I

Background Review
Chapter 1

Related Work

This research is multidisciplinary and explores several domains, including motion capture and analysis, biomechanics, signal processing and music. This chapter follows a top-down approach to outline different related works. In Section 1.1, we first focus on motion capture in general, then discuss papers that address specific aspects of the present project. In Section 1.2, we first focus on analysis of movements in general, and then on hand movements and musicians expert gestures.

1.1 Motion Capture

Motion capture appeared together with the development of instantaneous photography and cinematography at the end of the 19th century, making possible movements recording. Before the rise of this technology, it was difficult to perceive and measure complex movements. At that time, Etienne-Jules Marey invented cyclography, the ancestor of recent optical motion capture methods. According to that technique, a patient making a periodic movement, and wearing a black suit with narrow white tapes placed along each limb of his body, is captured with several photographic exposures on a single plate. This produces overlapping pictures allowing direct representation of motion on a single image, and analyze it with accuracy, thanks to the white tape. Cyclography was further developed into kymocyclography by Nikolai Bernstein in 1927, using electric bulbs as markers instead of tape, and a slowly and evenly moving photographic film instead of a single plate. The bulbs captured on the photographic film drew wavelike curves, easy to decipher [2].

The next step in the development of optical motion capture technology was stereoscopic recording of movements, allowing recording of an object with three spatial coordinates. This was achieved by recording the same scene from different points of observation. Details on the history of motion capture systems can be found in [2], [3] and [4].

Recently, different motion capture systems have been developed to allow very accurate measures. There are optical and non-optical systems. Optical systems, which use light reflections to capture objects, are widely used in research. Other systems exist, like mechanical systems, using exoskeletons, or systems using magnetism. All these systems can be divided in two categories: intrusive and non-intrusive systems. Intrusive systems use elements fixed on the object to be captured, like an exoskeleton, or optical markers. Non-intrusive systems do not need placement of intrusive elements on the target. They include cameras such as the Kinect from Microsoft, and OpenStage from Organic Motion. These systems, based on cameras, have a significant advantage as they are not intrusive, and hence allow freer target movements. Kinect allows the extraction of a 3D-map from a single 3D depth infrared sensor, while OpenStage is a multi-camera system, using shape-from-silhouette construction to extract a visual hull of a body. However, the accuracy of these markerless systems is still below that of intrusive systems [5, 6]. These systems are therefore more adapted to less demanding applications.

2. OpenStage: http://www.organicmotion.com/open-stage-2dot4-release/
In the present research, the motion capture system is used to analyze movements of the hand. Hence, the system must be able to capture complex movements of a small object, i.e. the hand, with a good accuracy. Therefore, an optical intrusive system is more appropriate.

The system used in this research is an intrusive optical motion capture system manufactured by Qualisys\(^3\), provided by the CIRMRT at McGill University. This system has the advantage of being very modular, as cameras and markers can be placed anywhere, in large and various amounts. It is also very accurate, compared to markerless systems. It is described in more details in Section 3.1.

1.1.1 Hand Motion Capture

For the particular case of the motion capture of the hand, several methods are proposed in the literature, using different technologies. Each method has its advantages and drawbacks in terms of cost and effectiveness.

A common method is the use of a sensing-glove. In [7], a glove packed with 20 sensors is used for acquisition of every joint angles (the Humanglove from Humanware\(^4\), see Fig 1.1 (a)). Technologies from CyberGlove\(^5\) are used in [8] and [9] for analysis of piano playing. The sensing-glove allows the capture of movements that could not be captured with camera-based technology, because of occultation problems\(^6\). However, their accuracy is far below that of optical systems [4]. They are also very invasive for the user. It is therefore not very suited for the purpose of the present research.

In [10], a color glove is proposed for hand tracking (see Fig 1.1 (b)). A simple glove is composed of different recognizable color patterns. The use of a simple camera allows recognition of the hand pose, relying on a hand pose database and nearest-neighbor search algorithm. This method allows fast and easy virtual reconstruction of the hand. This system is inexpensive, using commodity components. However, this simple method is limited to simple gestures in front of a camera and approximations restricted to the hand pose database. It is thus appropriate for basic hand animation or user interface control, but not for our specific application.

MacRitchie and Bailey (2013) [11] proposed an original low-cost method for pianist’s hands motion tracking. The proposed system is a monocular camera setup placed above the keyboard used to track markers directly painted on hands with UV passive paint (see Fig 1.1 (c and d)). This method allows efficient capture of 2D images, but the depth relies on estimations based on these 2D images. The precision is hence far poorer than with multi-camera systems allowing real 3D motion capture. Moreover, the occultation problem is greater with a single-camera system, degrading even more the capture quality.

Despite all these inventive methods for hand motion tracking, the systems mainly used in recent research are intrusive optical motion capture systems like Qualisys or Vicon\(^7\) [4]. Indeed, despite their price, these systems allow more flexibility and accuracy in measurements.

1.1.2 Hand Motion Capture using an Intrusive Optical System

The elaboration of an efficient motion capture setup requires an optimization step, so that the setup is suited for the specific capture. The system used in this research is an intrusive optical system, using several infrared cameras, and passive retro-reflective markers, as described in Section 3.1. This type of motion capture has the advantage to be very flexible, as markers of various sizes and shapes can be placed anywhere on a body, or on the objects to be capture. For an intrusive optical motion

---

3. Qualisys: www.qualisys.com
4. Humanware : http://www.hmw.it/hmw_e.html
6. Occultation problems occur when a target object (e.g. a hand) is hidden to the camera by another object (e.g. the other hand).
capture system, the optimization issue can be divided in two important points: the camera placement, and the marker model.

**Camera Placement**

The camera placement of optical motion capture systems has not been much explored in the literature, although it is not an obvious problem.

Olague and Mohr (1998) [12] developed an automated process to find an optimal placement of the cameras. This process is not specific to the study at hand, does not take account of possible occultation problems, and assumes fixed distances between the cameras and the targets. Chen and Davis (2000) [13] investigated camera placement of a motion capture system, considering camera resolution and occultation of the targets as constraints.

Both studies are very general, and do not take account of peculiarities of the specific scene and performance to be captured. In Chapter 4, we explain the development of an efficient setup for the specific capture of both of a pianist’s hands during a complex performance. Through this specific example, we also try to give a logical and general approach for the development of an efficient setup, extendable to more than this specific research.

**Marker Model**

The second point in the elaboration of an optical motion capture setup is the marker model, i.e. the choice of markers and their placement on the target. Their placement can indeed have a major influence on the quality of the capture. This point is well documented in the scientific literature, especially in a biomechanical and clinical context. For the particular case of the hand, Metcalf et al. (2008) [14] outlined several studies about placement of markers for different goals.

Miyata et al. (2004) [15] proposed a protocol for acquisition of complete and reliable data of the hand. They discussed the problem of closeness between markers on the hand, and the skin movement effects leading to inaccurate measures. The proposed model used a reduced number of markers (25...
exactly) to represent most Degrees Of Freedom (DOF) of the hand.

Su et al. (2005) [16] used a model with two markers on each finger joint, to take account of the skin movement effects. This model aimed at a better precision in measures of characteristics of patients’ hands.

The model proposed by Carpinella et al. (2006) [17] comprised 18 markers and allowed global capture of the position of the hand (see Fig. 1.2 left image). It was sufficient for the measure of Ranges Of Movement (ROM) of the hand in a clinical context.

Another model was proposed by Miyata et al. (2004) [15], using three markers on each joint to form a triangle (see Fig. 1.2 right image). It allowed determination of the exact orientation of each joint, and hence the precise capture of each DOF of the hand. This type of model can be useful in biomechanics, ergonomics, robotics, or advanced animation, when a strong hand model can be required. However, it is much more complicated to capture, due to the closeness between markers. In addition, it is also very cumbersome and can thus be very inconvenient for the patient, or performer in our specific application.

Metcalf et al. (2008) [14] proposed a new marker model composed of 26 markers, very similar to the one presented in [15]. Markers were placed on specific locations on the hand, particularly easy to identify by palpating the hand of the patient, and easy to apply. Metcalf et al. then proposed a calculation method to extract global features of the hand, based on planes and lines formed by the markers. The method allowed general characterization of the hand pose, by calculation of each joint angle.

Various models, simpler or similar to the model proposed by Metcalf et al. (2008) [14], were used in the context of the study regarding musicians’ expert gestures [18, 19, 20, 21, 22]. However, unlike in articles devoted to biomechanics, their choices were not discussed.

In the context of the present project, two other marker models are tested, and their advantages and drawbacks regarding our specific application are discussed in Chapter 4.

Figure 1.2: On the left: model comprising 18 markers for ROM measures [17]. On the right: model with three marker on each segment [15].

1.2 Motion Analysis

Together with studies and advances in motion capture, motion analysis has been explored. A pioneer in both fields was the Russian scientist Nikolai Bernstein (1896 - 1966), a key figure in the early study of biomechanics and motor control. Bernstein wrote several books explaining motor control and development, and coordination, published a long time after his death, because of the strict politics of the USSR during Bernstein’s living. Even today, his work is innovative and researchers in many fields may find interest in it, including biomechanics and motor behavior, neurology, sport and music education, physical therapy, etc. [23]
Recent research in motion analysis, regarding the entire body, is vast, and involves a wide range of research fields, including anthropology, biomechanics, neuropsychology, and motor control and behavior. More and more people try to make sense of how we move for applications like education, therapy, robotics, music, dance, sport, animation, games, HCI, and ergonomics. Burger and Toivainen (2013) [24] wrote a very general toolbox for analysis of motion capture data. This toolbox allows several types of analysis, including kinematic and kinetic analysis, time-series analysis, and analysis based on projections such as PCA. The toolbox has been originally developed for the study of musical gestures, but is applicable to any other motion analysis. The toolbox is compatible with several optical motion capture systems, including Qualisys. The toolbox is presented in Chapter 3, and most codes written for the present research are inspired from this toolbox.

1.2.1 Analysis of Musicians’ Expert Gestures

In this section, we focus on research devoted to the analysis of musicians’ motor control and behavior. Research in this context has been very active since the rise of motion capture technologies. Research on musical gestures, as well as sport gestures, can provide much information on motor control and motor learning. The analysis of these movements is indeed particularly interesting for two important reasons: they require great dexterity and motor coordination, and they follow an important formalism or rules, making the analysis better focused and hence more objective.

Musical gestures in general, and pianists’ expert gestures in particular, mainly involve arms, and particularly hands and fingers. The upper limbs of the body, and especially the hands, are very complex and exhibit numerous redundant DOF, justifying the relevance and interest of their study.

In 1930, Bernstein and Popova [3, 25] published a remarkable study on the "biodynamics of the piano strike". Their analysis and the method used to obtain data ("kymocyclography", developed earlier by Bernstein, see Section 1.1) were highly innovative at that time. They analyzed the dynamics of shoulder, elbow and wrist joints, during series of octave strikes performed by concert pianists. These series of octave strikes were executed with the right hand, at different tempi. This specific exercise allowed Bernstein and Popova to make strong interpretations about the "movement segmentation", a term widely used in recent research on motor control and behavior. They observed that, at a slower tempo, motion was segmented in well detached steps, while at a faster tempo, the movement became continuous. Moreover, the groups of muscles used, as well as the coupling between limbs, were different, depending on whether the movement was segmented or continuous. Although the aim of this analysis was a better understanding of pianists’ expert motor control and behavior, their findings about movement segmentation are widely extendable to the general study of motor control and behavior.

Many recent papers investigating hand control of musicians, and particularly pianists, are highly inspired from Bernstein researches. Indeed, these analyses are generally based on simple one-handed exercises executed at different tempi, and changes of dynamics are observed for specific parts of the upper limb.

Goebel and Palmer (2009) [19] studied the dynamics of the fingers according to the tempo. The exercise performed by pianists was a simple note progression and regression with the right hand, including each finger. The analysis showed that the finger trajectories changed with the tempo. This information is interesting in a general motor control context. It is also important in piano pedagogy, as it shows that the tempo used in the practice of an exercise is very important, as it can change considerably the movements used to achieve the exercise.

Furuya et al. (2011) [20] investigated distinct inter-joints coordination in piano playing, and showed how the nervous system of a skillful pianist coordinates the redundant DOF of the arm and hand for a better control, optimizing the muscles solicitation. This analysis was based on an exercise of alternate keystrokes of the thumb and pinkie (tremolo) at different tempi. It showed that at a faster tempo, an expert pianist made more use of the pronation-supination movements of the elbow, rather than
flexion-extension movements of the fingers. An electromyography analysis showed that this muscles
solicitation was more adapted to the tempo, for expert pianists, reducing co-activation and relaxing
muscles of the fingers. This analysis was again limited to the right hand, and to a specific gesture (the
tremolo technique).

Dalla Bella and Palmer (2011) [21] examined finger kinematics of different pianists, according to
the tempo and key velocity. The analysis was based on four skilled pianists, playing a simple melody
at different tempi. It showed that movement kinematics were related to a strategy aiming a tradeoff
between speed, timing accuracy and key velocity accuracy (related to the tone intensity). At a faster
tempo, finger height above the piano was generally greater. This increase of finger height led to higher
key velocities, but it was also correlated with a higher temporal accuracy. This analysis showed that
piano playing requires timing as well as spatial accuracy, both related to the skilled kinematics of a
pianist. A further analysis based on a neural network recognition process also showed that finger kine-
matics can be an indicator of personal signature of pianists. This signature linked to the movement is
also directly related to the resulting sound of the performance.

Goebel and Palmer (2013) [22] focused on pianists' finger and hand movements efficiency and tempo-
ral control in fast performance. They studied joints solicitations in a keystroke for different tempi. The
analysis showed that the metacarpophalangeal joints\(^8\) provide the main contribution to a keystroke.
On the contrary, the interphalangeal joints\(^9\) are opposite to the keystroke movement (finger extension).
Results showed that more skilled pianists made more efficient movements, avoiding opposite contribu-
tions to the keystroke movement, meaning less muscular solicitation. It also showed that more skilled
pianists, i.e. with more movement efficiency, also have a better temporal control in fast performance.

In the frame of the present research, two studies were achieved. The first one was in the same style
of approach as the previous ones. A simple exercise was executed by several pianists, at different tempi.
However, other constraints than the tempo were imposed to pianists. They were indeed asked to play
with different techniques, using more a group of muscles or another. This exercise aimed to show how
a pianist can play the same pattern differently. The ability to use and mix different redundant DOF
of the upper limb for the playing of a same pattern, would theoretically lead to an infinity of ways
of playing this pattern. The use of this feature could extend possibilities in the design of new digital
musical instruments. The mapping of these infinite ways of playing to the synthesis of a wide range of
new sounds could indeed provide a new dimension to the piano playing.

Additionally, this exercise was executed with both hands. All previous studies were limited to
one-handed exercises. A differentiation between both hands dynamics may thus be observed with this
analysis.

The second analysis is on the other hand highly different from the previous ones. It is quite an
innovation, as pianists both hands were analyzed on the game of real complex piano pieces, instead of
simple one-handed exercises. To do such, a highly efficient setup was required to capture both hands
on a whole piano piece. The acquired data were then used for PCA of each hand, for different pianists,
and for different piano pieces. Most previous studies concerning pianists hands analysis were limited to
time-series or kinematic analysis. PCA allows the extraction of more information on pianists’ expert
gestures, as well as information on the analyzed piano pieces.

In addition, PCA can provide another important information in the perspective of new digital
musical instruments design. It can indeed tell how pianists’ hand gestures are complex, and hence how
much data is needed to accurately represent these gestures as input of a new instrument.

---

8. The metacarpophalangeal joint is the joint between the palm and a finger. A background review of the hand
anatomy and biomechanics is given in Chapter 2.

9. The interphalangeal joint is the joint between two phalanges of a finger (see Chapter 2.)
Chapter 2

Hand Biomechanics

2.1 Introduction

This chapter briefly outlines the anatomy and the mechanics of the hand. A global understanding of the hand’s operation is important in order to conduct an adequate study, even in a non-medical context. This background review aims to explain the complexity of the hand motor control, depending at the same time on a muscular, an articular, and a neurological aspects, justifying why we can speak about expert gestures. This chapter can also provide a theoretical background to the reader, helping him fully understand every argumentation developed in this report.

The hand is a very sophisticated mechanism, allowing a huge number of movements. Among all the capabilities of the hand, the principal one is prehension. This ability is due to the possibility to oppose one finger, the thumb, to the others. With technology advances, the development of new tools has exploited a much wider range of hand movements. Indeed, though the hand was originally mainly used for grabbing things, it is now used for pressing a button, typing, plucking strings and playing chords on a guitar, or arpeggios on a piano.

2.2 Anatomy

The hand itself has a high dexterity, but it is also guided by an elaborate mechanism. Its steering system has seven DOF: three on the shoulder, two on the elbow, and two on the wrist. The hand proper is theoretically endowed with 23 DOF (see Fig. 2.1) [26]. As a practical matter, the hand is not as free, due to several causes, as we discuss in Section 2.4.

Fig. 2.1 shows the anatomy of the hand. The hand is attached to the arm by the radio-carpal and ulna-carpal joints, which form the wrist. The hand itself is composed of five metacarpals, linked to the wrist by carpometacarpal (CMC) joints. Each finger, linked to a metacarpal by a metacarpophalangeal (MCP) joint is composed of three phalanges, except for the thumb that has only two phalanges. The three phalanges are called proximal, middle and distal. The joints between phalanges are called proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints.

The thumb is a particular finger, endowed with three DOF on the CMC joint: flexion-extension, adduction-abduction and pronation-supination. These three DOF allow the opposition of the thumb to the palm, enabling the prehension task. The other four metacarpals are linked between them by inter-metacarpal junctions. The first three are fixed, but the pinkie is endowed with one DOF, allowing flexion of the CMC joint in a very short range, limited by the junction with the ring finger’s metacarpal. Flexion of the two mobile metacarpals (thumb and pinkie) allows a curving of the palm, and thus a rapprochement between the first and last fingers. This feature of the hand shows again that it is

---

1. This chapter is mainly inspired from Versier [26].
2. pronation-supination movement corresponds to the thumb rotating on itself, allowing opposition of the thumb fingerprint to the palm. Images are shown in Fig. 2.3 for flexion, extension, adduction, abduction of the fingers.
Figure 2.1: Anatomy of the hand. The numbers indicate the DOF for each joint of the hand. Adapted from [26]

designed mainly for prehension tasks. Moreover, the metacarpals are transversely and longitudinally curved, so that the palm is naturally hollow, also facilitating gripping tasks [26].

2.3 Muscles and Movements

The hand is controlled by many muscles, located in both the palm and forearm [26, 27] (see Fig. 2.2). The movements of the wrist are controlled by several muscles of the forearm. Each muscle has a specific function: extension, flexion, abduction or adduction. The other muscles in the forearm control flexion and extension movements of the fingers, including common and individual movements.

The muscles of the palm are less powerful, but allow more precise movements. These movements include flexion and extension, but also abduction and adduction movements of each finger separately. The general movements of the fingers are shown in Fig. 2.3.

According to Kapandji (1995) [28], there are three types of finger flexion, using different muscles groups. There is the "shutter flexion" bending only the MCP joint by using only intrinsic muscles, the "hook flexion" bending only the DIP joint by using only extrinsic muscles, and the "harmonious flexion" bending the three phalangeal joints, necessitating a balance between intrinsic and extrinsic muscles. Skilled piano playing requires the use of harmonious flexion that demands a proper motor control [29].

On piano, according to the style of playing, different DOF of the upper limb and hence different groups of muscles can be used. The wrist is in the middle of the chain of several redundant DOF, and plays an important role in the looseness or steadiness of playing. Detached and rapid sounds can be achieved by blocking the wrist and using flexion-extension movements of the elbow, while a loose wrist allows a precise and delicate playing [30].
2.4 Enslaving Effects

In principle, all these muscles should allow the hand to make any movement in accordance with its DOF. However, there are some limitations. For instance, it is extremely difficult for most people to move the pinkie without moving the ring finger unintentionally. These inter-dependencies between fingers, called enslaving effects [31], can be due to several causes:

1. inter-tendinous connections between the tendons of the fingers induce inter-dependencies between muscles of the fingers 2 to 5 (see left image Fig. 2.2) [27];

2. some fibers from one muscle can be attached to the tendon of another muscle. This results in a coactivation of the tendons [31, 32, 33];

3. the neuronal zones in the motor cortex that control muscles of the hands overlap. It can be helpful for coordination of the fingers in movements like prehension. The control of the prehension task thus becomes easier, by reducing the number of DOF to control [33, 34, 35].

Such enslaving effects can be useful for the prehension task, by coordinating movements, and simplifying the gesture. However, it also reduces the independence between fingers, and can have adverse
effects in several activities like playing the piano [33].

The third cause though, i.e. the neurological one, can be reduced by training. This is possible, thanks to the neuroplasticity of the human brain. Unlike the mechanical components like tendons and muscle fibers connections, synapses\textsuperscript{3} can be reorganized in the motor cortex after intense or repetitive training, in order to increase the efficiency of the neuronal activity for the exercise trained [36].

2.5 Conclusion

This chapter briefly presented the operation of the hand, and provided the necessary background and vocabulary to apprehend this report. The hand is a highly complex piece of technology, tailored by the human evolution, especially for its most important task, the prehension task. However, the same features that facilitate the prehension task, also limit the scope of movements that can be achieved with the hand. These limitations are due to three different enslaving effects: muscular coactivation, articular junctions, and neuronal zones overlapping.

Nonetheless, intensive training can force a rearrangement of neuronal zones, to adapt the motor control of the hand to the trained movements. This functionality of the brain is called neuroplasticity, and allows pianists to improve their motor control of the hand, in order to perform more complex and precise gestures, and master hard piano pieces.

PCA, used in Chapter 7 can provide a profile of complexity of the movements achieved by pianists during a performance. Results seemed to show that a great level of expertise in piano exhibits more complexity in the measured piano playing gestures.

\textsuperscript{3} connections between neurons
Chapter 3

Materials

An understanding of the working of the tools, toolboxes, and algorithms used in this research, may help the reader towards a better apprehension, and comprehension of the whole project. This chapter presents the different tools used during this research. Section 3.1 focuses on the motion capture system used in this research: Qualisys. Section 3.2 briefly presents the MIDI protocol (Musical Instrument Digital Interface), used to capture other useful data in the analysis of the pianist’s expert gestures. Section 3.3 presents two toolboxes used to import motion and MIDI data into the MATLAB environment, and to process and analyze them. The last section, Section 3.4 presents then the algorithm of PCA. PCA is a very powerful tool, used in the frame of this research, to analyze the movements of the pianist’s hands.

3.1 Motion Capture System

The motion capture system is a tool used for the acquisition of motion data, i.e. the movements of the target to capture. This technology has been widely used in several domains, like animation, biomechanics, music, etc. Some examples have been provided in Chapter 1. As explained in Chapter 1, the system used in this research is an optical intrusive system, manufactured by Qualisys\(^1\).

Qualisys is a marker-based motion capture system, using both active or passive markers. Active means that the marker emits light itself. On the contrary, a passive marker is a small bowl with retro-reflective tape on it, which only reflects the available light. In the setup developed in this project, only passive markers were used.

3.1.1 Passive Markers

There are several types of markers, differing by their size, and by their shape. Each marker consists of a small sphere, or hemisphere, covered with reflective tape (see Fig. 3.1).

![Figure 3.1: Passive markers types [37]](image)

Table 3.1 lists the different types of markers provided by the CIRMMT which were used for this project. The choice of the shape and size of the markers depends on several factors, like the distance to

---

1. Qualisys: http://www.qualisys.com/
the cameras, the size and shape of the object to capture, or the type of placement. Big markers can be seen in a wide range but are cumbersome and difficult to place on small objects. On the opposite, tiny hemispherical markers can be placed everywhere, like on human lips, fingers, or small mechanisms, but require an adapted acquisition system to be captured [37].

Table 3.1: Types of markers

<table>
<thead>
<tr>
<th>Size (diameter)</th>
<th>Shape</th>
<th>Flat basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>4mm</td>
<td>Hemispherical</td>
<td>No</td>
</tr>
<tr>
<td>7mm</td>
<td>Hemispherical</td>
<td>No</td>
</tr>
<tr>
<td>7mm</td>
<td>Spherical</td>
<td>No</td>
</tr>
<tr>
<td>7mm</td>
<td>Spherical</td>
<td>Yes</td>
</tr>
<tr>
<td>12mm</td>
<td>Spherical</td>
<td>Yes</td>
</tr>
</tbody>
</table>

3.1.2 Acquisition System

The acquisition system is a cameras network, linked to a computer equipped with the Qualisys software, Qualisys Track Manager®. Each camera is equipped with an infrared strobe. The infrared spectrum has been chosen not to disturb the user. Fig. 3.2 shows the acquisition system.

The cameras are plugged in series, by bundled cables including power and communication cables. The cameras communicate with each other and with the computer, thanks to the Ethernet protocol. This communication channel is used for configuration and synchronization of the cameras, and for data transmission. Each camera is provided with an integrated system that allows options configuration, and pre-processing of the captured images.

3.1.3 Data Acquisition

When a marker is seen by one camera (in 2D), the location of the possible positions of the marker is a ray (a half-line), starting at the camera’s lens and crossing the marker. If two cameras see the marker, its actual position can thus be found by calculating the intersection between two rays. According to the resolution of the cameras, the result will be more or less accurate, and can be improved by calculating the intersection with rays of more than two cameras.

To capture precise movements, the acquisition system must capture markers positions at a high-speed rate\(^2\). As positions are determined by calculations based on several cameras data, the cameras must be perfectly synchronized to capture the markers exactly at the same time.

Each frame is captured by emitting a flash with the strobe of the camera, and capturing the light reflected by the environment and the retro-reflective markers. Each camera captures an intensity

\(^2\) The cameras used in this study were Oqus Cameras Series 4. The maximum frame rate of these cameras in marker capture mode is 480 FPS (Frames Per Second).
image and calculates the positions of the markers in 2D, by applying a threshold on the intensity of
the captured light. If the threshold is adequately defined, only peaks of intensity corresponding to the
rays reflected by the retro-reflective markers will be kept. The camera then calculates for each peak
its center, and two axes of an ellipse, corresponding to the shape and width of the peak. Each camera
sends then these 2D-data to the computer through the communication channel, and the 3D-positions
are processed, assuming known positions of every camera.

3.1.4 Calibration

The preceding section assumes that the computer is already aware of the positions of every camera.
Otherwise the positions of the markers cannot be obtained. To know the positions of the cameras
relative to each other, a calibration step must first be achieved. This is done by using references, or
calibration tools (see Fig. 3.3).

![Figure 3.3: Calibration tools [37]](image)

Fig. 3.3 shows two instruments used to calibrate the cameras installation. The L-shaped instrument
is the reference structure. It is placed in the middle of the scene to capture before the calibration step.
The origin of the coordinate system of the scene is defined on the angle of this structure, and the axes
are formed by its two perpendicular segments, and their vector product. The T-shaped instrument
is the calibration wand. A calibration capture must be made, while waving the wand in the capture
area, and covering the whole measurement volume. These two instruments have predefined dimensions,
and hence predefined distances between the markers, which are known by the Qualisys software. A
capture process of these markers allows accurate extraction of the position and orientation of each
camera relative to the reference structure, and to scale properly each axis of the coordinate system in
the measurement volume [38].

3.1.5 Trajectories Computation

The last step of a capture session is to identify markers, and calculate their most probable trajecto-
ries by comparing each frame to the previous and next ones. This step is done by the software Qualisys
Track Manager. This program also allows filling gaps in trajectories by polynomial interpolation, when
markers disappear for a short period of time.

3.2 Keyboard - MIDI Protocol

Motion capture data are not the only useful data for gesture analysis. As well as the motion cap-
ture acquisition system, the digital piano, or electronic keyboard, is a bundle of sensors itself. As a
matter of fact, every key can be considered as a velocity sensor. The keyboard used for this research
was a Yamaha CP300. Like most current keyboards, it allows collection of data from the keys via a
USB-port. The communication protocol used to transmit data messages between the keyboard and a
computer is the MIDI protocol. During the motion capture sessions of this research, data messages were recorded on the computer thanks to the freeware Anvil Studio\textsuperscript{3}. Anvil Studio is a MIDI acquisition program, as well as a MIDI editor and a music composition program. It allows saving the messages recorded as a MIDI file.

A MIDI message is basically composed of three bytes: the first byte gives a MIDI channel (1-16), and a command or event code. For instance the "Note On" or "Note Off" events are the codes sent when a key is pressed or released. The second byte gives the note number, i.e. the index number of the key that is pressed or released (0-127). As a classical keyboard is composed of 88 notes (from A0 to C8), the actual range of this byte is 21-108 (see Fig. 3.4). The third byte gives the note velocity (0-127), telling the speed at which the key was pressed or released. Each key press thus generates two MIDI messages. A message is sent including the "Note On" event when the key is pressed, and another message including the "Note Off" event is sent when the key is released\textsuperscript{39}.

![Figure 3.4: 88-notes classical keyboard - Note names and MIDI numbers.](Image)

The MIDI editor allows saving every MIDI messages sent by the keyboard, along with a timestamp for each message, corresponding to the time when the message was received. These messages, along with their timestamps, can then be saved in a "*.mid" file.

In the context of this research, the MIDI data were used mostly as landmarks for the motion data, for renditions comparison and for data segmentation.

### 3.3 Matlab Toolboxes

To perform analysis on the data captured with Qualisys and recorded with Anvil Studio, The program MATLAB\textsuperscript{®} was used. Two toolboxes were added to the program: the MOCAP Toolbox\textsuperscript{24}, and the MIDI Toolbox\textsuperscript{40}. These toolboxes respectively allow importation of motion data and MIDI data into the MATLAB environment, and consist of many functions allowing handling and analysis of these data.

#### 3.3.1 MOCAP Toolbox

The MOCAP Toolbox\textsuperscript{4}, developed at University of Jyväskylä, Finland, allows importation of motion data recorded with Qualisys into the MATLAB environment, as a structure, containing the data and all kinds of useful information, like the framerate of the capture, the number of markers captured, etc. Many functions are included in the toolbox, and allow data manipulation for different purposes. These functions allow data editing and transformation, visualization, and various kinds of analysis. The analysis functions are divided in several categories: kinematic, kinetic and time-series analysis. Another particular category of functions are the projection functions, allowing for instance to perform a projection on the first Principal Components (PC) of the movements\textsuperscript{24}. The functions used for this study were mainly employed for edition, transformation and visualization of data, as well as for PC projection.

\textsuperscript{3} Anvil Studio: http://www.anvilstudio.com/.

\textsuperscript{4} The MOCAP Toolbox is available at: https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mocaptoolbox.
3.3.2 MIDI Toolbox

The MIDI Toolbox\textsuperscript{5}\textsuperscript{[40]} was also developed at University of Jyväskylä. It provides codes to import a MIDI file into MATLAB, as a matrix. Each row of the matrix corresponds to a note pressed during recording. The rows are sorted according to the timestamps of every key press. The first row corresponds to the first note played by the pianist, and the last row to the last note played. The matrix has seven columns containing various information, and is organized as presented in Table 3.2.

<table>
<thead>
<tr>
<th>Onset (beats)</th>
<th>Duration (beats)</th>
<th>MIDI channel</th>
<th>MIDI pitch</th>
<th>Velocity</th>
<th>Onset (s)</th>
<th>Duration (s)</th>
</tr>
</thead>
</table>

The onset is the timestamp of the beginning of the keypress, that is of the "Note On" event (see Section 3.2). The duration is the difference between the timestamps of the "Note Off" and the "Note On" event. The MIDI channel is used for multitrack MIDI files, including for example several melodies played by different instruments, and is therefore not important for this project. The MIDI pitch indicates the note number, and the velocity gives the speed of the keystroke and thus the loudness of the note, on a scale of 0 to 127. The sixth and seventh columns give the same information as the first two, in a different unit system. The ratio between these columns is the tempo (beats per minute) of the MIDI file. The tempo is a basic parameter of a MIDI file. Its default value is 100 or 120 beats per minute. A modification of this parameter allows changing the speed of the whole MIDI score, recorded in beats.

The MIDI Toolbox was used in this research to import MIDI files into MATLAB, and write new MIDI files after processing of this matrix.

3.4 Principal Component Analysis

PCA is a mathematical tool used in several domains, like signal processing, or statistics. According to the application field, the utility of PCA can be different. In the scope of this research, the PCA algorithm was used for two purposes: analysis of the complexity of the pianist’s expert gestures, and feature extraction of the pianist’s movements. If the aim is multi-purpose, the principle is the same: a data representation with a lower dimensionality.

When a data set is represented with many components, or dimensions, there are often high correlations between these components. As an example, if a finger is represented with a marker on the first phalanx, and another on the second one, the correlations between the markers are high as their movements are almost identical. It means that the second marker does not provide much additional information on the movements of the finger.

The PCA process consists in a linear transformation of the data representation system (or vector base), which leads to the best representation of these data, according to the least square criterion [41]. As a matter of fact, PCA leads to a data representation with decorrelated components. It can also be said that PCA tends to diagonalize the covariance matrix or correlation matrix of the data.

Let $X$ be a $n - by - d$ matrix, representing a data sequence of length $n$, in a $d$-coordinate system. The covariance matrix is given by\textsuperscript{6}:

$$C = X^T X$$

\textsuperscript{5} The MIDI Toolbox is available at: \url{https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/midtoolbox}. Nevertheless, a modified version was used in this project, as the MIDI Toolbox is limited to 32-bit architecture computers. The modified version is available at: \url{http://www.ee.columbia.edu/~csmit/matlab_midi.html}.

\textsuperscript{6} The data are considered to be centered on the center of gravity before the process, in order to avoid the impact of their offset on the process. It may as well be appropriate to reduce data before the process. It allows giving the same weight to every component in the process.
where $C$ is a $d \times d$ symmetric square matrix. If this matrix is diagonal, it means that the dimensions of $X$ are not correlated. The information given by each dimension $i$ is provided by its own variance $x_i \cdot x_i$ (the diagonal elements of $C$), as it is an indicator of the dispersion of data on that axis. A large dispersion in that dimension means that this component is substantial and provides much information on the data set. On the opposite, if this component has the same value for all data it does not provide any information.

Let $\lambda_i, (i = 1, \ldots, d)$ be the eigenvalues, and $v_i, (i = 1, \ldots, d)$ the eigenvectors (of dimension $d \times 1$) of $C$. As $C$ is symmetric, its diagonalization is given by:

$$
P = (v_1 \ldots v_d), \quad (3.2)
$$

$$
L = \begin{pmatrix} 
\lambda_1 & & \\
& \ddots & \\
& & \lambda_d 
\end{pmatrix}, \quad (3.3)
$$

$$
\text{diag}(C) = P^T C P = L \quad (3.4)
$$

where $P$ is a $d \times d$ orthonormal matrix (or vector base) formed by the eigenvectors of $C$, and $L$ a $d \times d$ diagonal matrix, formed by the eigenvalues, and result of the diagonalization of $C$ ($\text{diag}(C)$ means diagonalization of $C$).

As $C$ is symmetric, its eigenvectors form an orthonormal base, represented by the matrix $P$. If we represent $X$ in that orthonormal base:

$$
X_p = XP, \quad (3.5)
$$

It can be shown that the covariance matrix of this new matrix $X_p$ is diagonal:

$$
X_p^T X_p = (XP)^T (XP) = P^T X^T X P = P^T C P = L \quad (3.6)
$$

The dimensions of $X_p$ are thus decorrelated, and the variance of data on each axis (eigenvector) is given by the corresponding eigenvalue of $C$.

PCA uses this property to extract the PC of the data. The most representative components of $X_p$ are those whose dispersion or variance is higher, and hence whose corresponding eigenvalue in the matrix $L$ is higher.

As the whole set of components provides 100% information on the data, the eigenvalues of $C$ indicate the weight of information given by each component. If most information is given by a few components, it is possible to extract an approximate representation of the data with a lower dimensionality. For more information, an illustrative example is provided in Appendix A.

In the context of this research, PCA was used for two purposes:

- analysis of the complexity and diversity of the pianist’s expert gestures. The computed eigenvalues can indicate a complexity, or diversity profile of the gestures performed by the pianist;
- feature extraction of the pianist’s movements. In the scope of gesture recognition, the extraction of consistent features is very important. Features too numerous and not consistent would lead to a poor performance of the recognition process. PCA allows extraction of the most representative components of the gestures, as features for the recognition process.
Part II

Data Collection
Chapter 4

Motion Capture Setup

4.1 Introduction

This chapter presents and discusses the development of a setup optimized for the study of the pianist’s expert gestures. Fig. 4.1 shows a block diagram of the general setup. The input data are the scene, including the markers placed on the body and the keyboard, as well as the MIDI messages generated by keypresses on the keyboard. The markers are captured by the cameras of the motion capture system and the marker trajectories are then calculated. The MIDI messages are recorded by the MIDI acquisition software and written in a MIDI file. The marker trajectories and MIDI data are then imported into MATLAB and a synchronization process is applied.

![Figure 4.1: Motion Capture Setup](image)

In motion analysis, the elaboration of an efficient motion capture setup is very important to get adequate data to analyze. Two issues must be approached in conjunction: the marker setup, and the camera setup.

As explained in Chapter 1, a setup is generally specific to the study at hand. Still, it is possible to have a general qualitative understanding of the issue, based on logical reasoning, and this chapter aims to develop this reasoning with the explanation of the particular case of this project. The general approach applied on this problem, may then be reproducible on related but different projects.

Our methods are presented in Section 4.2. They focus especially on the capture of the hands, as it is by far the most complicated, and the most interesting part to capture and analyze. Different setups of markers and cameras are presented. The MIDI acquisition and synchronization methods are also explained. Results are then shown and discussed in Section 4.3.
4.2 Methods

The goal of this part of the project is the establishment of a motion capture setup that allows the capture of both hands of the pianist, for any rendition. A marker-based motion capture setup implies two main issues:

- the marker setup: the type of markers and their positions, accounting for the specific study, and for a good representation of the objects to capture;
- the camera setup: the number of cameras, and their positions according to the objects to capture, like the hands, and obstructive objects, such as a piano or the rest of the body.

Depending on the application, different marker setups can be adequate. Some models can require only a few markers on the hand, placed to acquire only some specific data such as wrists angles. On the contrary, other studies aiming for very accurate measures, imply several markers on each joint, to compensate for the effects of the skin movements on the skeleton, and hence deviation of the marker (see Chapter 1). These examples show that the marker setup is very specific to the study.

In the context of this project, the issues of the marker setup are multiple:
1. in order to fully capture the movement of the hand, every DOF of the hand must be represented;
2. the setup must be as little restrictive as possible for the pianist, so that he can play freely;
3. the markers must be as visible as possible, by at least two cameras at anytime.

In order to satisfy all these constraints, different compromises must be made in the setting up. Different setups of the markers have been tested, trying to fit the constraints. Section 4.2.1 discusses these constraints, by comparing two different models.

The issue of the camera setup is tightly linked to the marker issue. If the markers are hard to see, the camera installation must be adapted, and must consider every extreme case in the hands movements. These cases are discussed in Section 4.2.2.

4.2.1 Marker Model

According to the concerned study, the marker setup can be very different. In this project, several setups, or models, have been tested, with among them slight differences concerning only the type or position of some markers. The present section focuses on two models, from which other models have been derived.

Fig. 4.2 shows two different models of the markers on the hand. Each one has several advantages and drawbacks. In a marker model, two features are important:

- the placement of the markers. The position of each marker must fit with the study. A non-pertinent marker gives useless data. Moreover, every marker is noise for another marker, hence useless markers must be avoided. Two markers too close can lead to confusion, if the cameras are not precise enough to discern them. It can also be source of occultation, or swapping of the calculated trajectories between markers;

- the type of each marker. According to its position, a marker must have an adapted shape and size. If the marker is in a difficult position to see, it must be big enough to be captured more easily by at least two cameras. On the other hand, if it is near another marker, it cannot be too big, or it will hide the other markers. Moreover, markers in delicate positions, like on distal phalanges, must be tiny, or they will touch the keyboard when the finger is playing. Especially on the thumb, whose inclination (roll) is different from that of the other fingers, the markers must be tiny, not to restrict the pianist’s movements. Too big markers on the fingers can also be
a problem in hand or finger crossover techniques.

These two features are presented below for the two models (Fig. 4.2). As seen in Chapter 2 - Hand Biomechanics, the hand is a complex mechanism, and can be fully captured only by placing many markers on it, considering every DOF. The first model, called bone-based marker model, on the left image, is composed of 21 markers on the hand. The second model, joint-based marker model, is composed of 24 markers.

Figure 4.2: On the left: bone-based marker model. Markers are placed on the bones so the position of every bone can be measured at every instant. On the right: joint-based marker model. A marker is placed on each joint of the hand. This model allows direct calculation of each articulation angle.

Markers positions

On the first model (see Fig. 4.2, left image), markers are placed in the positions of the bones. This model allows indicating the position of most bones of the hand with a minimum number of markers. The second model, on the opposite, is joint-based (see Fig. 4.2, right image). Markers are placed on the joints, and the segment between each markers-pair corresponds to a bone. This configuration of the markers allows measurement of the flexion of each joint, by calculating the angle between the two corresponding segments. The second model can thus give a more complete representation of the hand, graphically and kinematically. That is the reason why similar joint-based models were used in several studies, in the context of animation [42], medicine [14] and analysis of musicians' expert gestures [22].

The joint-based model fits better to the first issue of the marker setup, i.e. representation of the hand. Nevertheless, it is not the case for the two other issues. Several markers are more restrictive and difficult to track. Indeed, the markers are really close, especially on the DIP joints and on the fingertips. However, these constraints can be reduced by adapting the markers types and the camera setup.

Markers types

As seen in Section 3.1, there are several types of markers, differing by their size, and by their shape. The type of each marker must be chosen according to its position, to improve the capture performance or the comfort of the pianist.
Size To better understand the importance of the size of the markers, we must consider the environment of the capture scene. Many cameras emit infrared light in the whole room. This light is reflected by the ground, walls, and more particularly by the objects at the center of the scene, like the keyboard, possibly sheets played by the pianist, and the body of the pianist. All this undesirable infrared light is seen by the cameras and is therefore noise for the motion capture system. The process of capturing markers is done by using a threshold on the infrared light seen by the cameras, and seeking peaks corresponding to the beams reflected by the markers. The intensity of the light reflected by a marker is roughly proportional to the surface of reflection and hence to the square of the radius of the marker. Bigger markers will induce bigger peaks in the camera’s field of vision, and will thus be more easily recognized. A too small marker can disappear from the capture if the peak reflected by the marker is beneath the threshold defined for the capture. Fig. 4.3 shows an intensity view of a camera placed in front of the piano. Several sheets are laying on the piano, along with the calibration tools. The color scale (from black to red) indicates the intensity of the infrared light seen by the camera. The white keys of the piano as well as the sheets are colored blue, as they reflect a lot of infrared light emitted by all the cameras. The markers of the calibration tools are colored red, meaning they reflect a lot more infrared than the rest of the scene. These markers are 12 mm spherical markers. A proper threshold on the image allows capture of the markers only and rejection of the noise.

![Intensity view of a camera](image)

Figure 4.3: Intensity view of a camera.

Markers shape It can be observed on the two models that, when the marker position is near the edges of the fingers, smaller markers have been chosen. For the joint-based model especially, we can see in Fig. 4.2 that 4 mm hemispherical markers have been placed on the tips, to ensure the pianist’s comfort. The cost of this choice is that the tips of the fingers are more difficult to capture, and the camera setup must be adapted, as developed in Section 4.2.2.

As their position is away from the tip, the markers are less uncomfortable for the pianist and can thus be bigger and more visible. On the last phalanx for the bone-based model, and on DIP joints for the joint-based model, 7 mm markers without a flat basis have been used, as a basis might be larger than the finger and hence disturb the pianist. For the rest of the hand, 7 mm spherical markers with a base have been used. Bigger markers would have been too cumbersome, and too close, inducing confusion, as the two infrared peaks would be seen as one bigger and broader peak by the cameras.

4.2.2 Camera setup

Another important issue of the motion capture installation is the cameras setup. An efficient camera setup must respect important rules. The first one is the following:
Rule 1. Each marker must be seen by at least two cameras at any moment of the capture period, so that its position can be computed for the whole period. The setting up must be achieved considering that rule, and the fact that the movements to capture, i.e. movements of the body and especially the hands of the pianist, are highly variable and complex. It means that the setup must be capable of seeing markers in many different positions, with at least two cameras at anytime.

Another issue to consider is the fact that a camera is always noise for another camera, as it generates infrared light that can be involuntarily captured by another camera, directly or through reflections. Another rule can thus be set:

Rule 2. Optimized positions and orientations of the cameras are those that minimize the noise generated by the other cameras. An obvious consequence of that rule is that two cameras cannot be directly in front of each other. It would induce huge noise, completely disturbing the intensity images of the two cameras, and hence the markers extraction.

In the context of this study, the cameras must be focused on the keyboard and hands of the piano player. They must be placed in adequate points of view so that each marker is seen by at least two cameras at any time, considering all the possible occultations of a marker, due to movements of the pianist or to the marker setup. The list below summarizes several causes of occultation, due to the pianist’s movements:

- a finger under another (finger crossover technique). A very frequent technique is the thumb under that allows a hand to make arpeggios covering the entire keyboard. It is done by passing other fingers above the thumb to play long sequences of notes on a large scale at a fast rate;
- a hand under another (hand crossover technique);
- the player’s body or head between the camera and the hands;

In addition to these possible occultations, we must consider all the possible orientations of the hands during a rendition. From an observation point of view, the hand can be thought of as a combination of observation planes. Fig. 4.4 shows a very simple example of observation planes of the hand on a piano. One plane corresponds to the back of the hand, the four long fingers roughly correspond to another plane, and the third plane corresponds to the thumb, whose orientation is very different from that of the other fingers. According to the variability of the pianist’s movements, the amount and orientations of the observation planes can vary in a wide range.

The best viewpoint for a camera from which to capture a plane is in the direction perpendicular to that plane. When oriented in this direction, a camera can clearly observe and distinguish every marker located in that plane. At the other extreme, if the camera is pointing in parallel to the plane, the markers located on that plane might be indistinguishable from each other by the camera. Furthermore, if the angle between the observation line of the camera and the normal vector is above 90°, every marker is hidden to the camera.

During a performance, both hands of the pianist can travel a long distance, cover a large volume, and the normal vector of each plane can point in every direction. Especially in the case of the thumb, the direction of the normal vector can change in a wide range. The thumb is indeed a particular finger, with a great mobility, and whose observation planes differ a lot from that the others. As a matter of fact, the thumb is the most complicated limb to capture.

The distance of each camera relative to the scene to be captured is also important. To efficiently capture small markers, the camera must be close enough, so that the peak of light reflected by a marker is high enough to be distinguished from noise. On the other hand, to be able to capture markers in a wide range, the camera must be far enough from the scene (according to its field of view). As the hands are likely to move along the entire keyboard, an adequate distance would be the one that allows
the camera to see exactly the full keyboard, and the hands waving above it.

The camera setup must be achieved taking the above considerations into account. First, we must consider every possible orientation of each normal vector during a rendition. With respect to Rule 1, at least two cameras must correctly see the plane, and thus the markers, for each of these orientations. Furthermore, the cameras must not be in front of each other, so as to minimize disturbances.

Considering the possible orientations of the planes, and the adequate distance of cameras relative to the keyboard, the location that the cameras should cover is roughly a hemisphere above the keyboard, centered on it, and with a radius equal to the distance necessary to see the whole keyboard. From this hemisphere, an area must be subtracted as it is behind the pianist’s body and hence a camera placed there would hardly see any marker placed on the hands.

Considering the first rule that two cameras must see each marker at every point in time, the angles between the observation lines of the cameras must be reasonably small, so that the two cameras are simultaneously close to the normal of an observation plane. As the cameras focus on the keyboard, and considering that they must not be in front of each other, they must be at least slightly higher than the keyboard and point downwards. To avoid indirect disturbances due to reflections on the keyboard from one camera to another, symmetry should be avoided.

Fig. 4.5 shows a setup based on the above considerations with a minimum number of cameras. This setup consists of six cameras. The five first cameras form an approximately regular pentagon, around the piano. The sixth camera is placed directly above the keyboard, except for a small horizontal shift to avoid occultation due to the pianist’s head. A photo of the setup and 2D marker-views of each camera of the setup can be seen in Appendix B.1.

A connection with the observation planes on Fig. 4.4 can be made. Camera #1 is in front of plane #2, corresponding to the long fingers. Cameras #2 and #3 can complete Rule #1 for this plane, and for both hands. Camera #6 faces plane #1, corresponding to the back of the hand. Every other camera can assist in completing Rule #1 for both hands as they all roughly see the hand’s back plane (see 2D-views, Appendix B.1). Pairs of cameras #2-4 and #3-5 allow observation of plane #3 corresponding to the thumb, for each hand. Their positions allow them to see the thumb even during a thumb under technique, when the thumb is below the rest of the hand (an example can be seen in Appendix B.1).

Moreover, when the pianist rises a hand up in the air while playing, the normal vector of plane #1 can sometimes point behind the pianist, in some extreme case. Pair of cameras #4-5 allows partial capture of plane #1 in this case.
Fig. 4.6 shows another possible camera setup, based on the previous considerations. This setup consists of twelve cameras, approximately located on a hemisphere centered at the mid-point of the keyboard. This setup requires many more cameras, but allows the capture of more extreme cases than the 6-camera setup. The first six cameras have roughly the same positions and roles as for the 6-camera setup, but cameras #2 and #3 have been doubled with cameras #7 and #8. On the 6-camera setup, the role of camera #2 (resp. #3) is double: to see the left thumb observation plane (resp. right thumb), and partially complete the first rule with camera #1, for the long fingers observation planes. The 12-camera setup allows splitting these roles: camera #2 is slightly shifted to better see the thumb, and camera #7 (or #2') correctly sees the long fingers (see Fig. 4.6). The analogy can be made for cameras #3 and #8 (#3'). Cameras #9 and #10 are more specific to the capture of the fingertips markers, on the joint-based marker model. These markers are the smallest ones, and are therefore difficult to capture. Moreover, the normal vectors corresponding to these markers can often be almost horizontal, justifying the low positions of cameras #9 and #10. Cameras #11 and #12 have been added to account for more extreme cases, but also to capture the rest of the body, like elbows, shoulders and the neck. This is why they are further away than the first ten cameras.

More generally, in this setup, more cameras have been placed in front of the sensitive regions, i.e. in front of all the fingers, where markers are smaller, closer, and hence more prone to occultation or confusion. A photo of the setup, and views of each camera, are shown in Appendix B.2.

The last step in the camera setup is the tuning of every camera. For each camera, the aperture and the focus can be manually tuned, and must be adapted to its position relative to the scene. The focal point must be near the center of the keyboard. The aperture must be tuned so that the depth of field corresponds to the keyboard length. The image must indeed be clear all along the keyboard, otherwise the markers will not be correctly seen. If the camera is in front of the keyboard (camera #1 for instance), the aperture can be wide, as the depth of field must not be large. In contrast, for cameras #2 and #3 which are on the sides of the keyboard, the depth of field must be larger in order to clearly see both sides of the keyboard, and the aperture must hence be narrowed.

4.2.3 MIDI Synchronization

Qualisys cameras and markers are used for acquisition of motion data. As mentioned in Chapter 3, other data can be useful for the analysis of the pianists’ performance. MIDI data help as a landmark, and as comparison benchmark for different renditions of the same pieces. It allows easy extraction of parts of a piece where pianists play exactly the same notes, making comparison more objective. Moreover, the quality of the participant’s rendition can be checked by listening to synthesized music from the MIDI file. Furthermore, we will see in Chapter 6 how MIDI data are used for automatic segmentation of a specific exercise.
MIDI data are recorded separately from motion data, and hence must be synchronized with the latter after the capture session. A very simple method is used in the setup, for this purpose: a reflective marker, so-called "reference marker", is placed on one piano key, and at the beginning of every capture session, the key is pressed. The first MIDI message thus corresponds to the first peak found on the z-axis (height) of the reference marker. The chosen reference note was A#7 (MIDI number = 106, see Fig. 3.4), the last black key of the piano (see Fig. 4.7).  

MIDI files and motion data files are loaded into MATLAB, and are then processed by a specific MATLAB function written to synchronize automatically every MIDI file to its corresponding motion data. 

---

1. During the calibration of the motion capture system, the reference structure (see Section 3.1.4) was placed on the keyboard, such that x-axis was along the keyboard (from left hand to right hand), y-axis pointed towards the body of the pianist, and z-axis was oriented upward.

2. This key was chosen because black keys are higher than white keys, and are therefore more practical for the placement of a marker; and because A#7 is the last key of the piano, hence less used than others.
data file. It is achieved by applying a threshold on the height of the reference marker: when the height is under the threshold, the reference key has been pressed. The MIDI data are then temporally shifted to fit with the keystress of the reference key, making the first MIDI message timestamp correspond with the peak occurring when the reference marker is lowered. A verification of the note number (106) of the first MIDI message is done to ensure that the message corresponds to the reference key. Fig. 4.8 shows the temporal shift of the MIDI data, making the first MIDI message timestamp correspond with the keystroke of the reference key. The green vertical lines correspond to the first MIDI message timestamp before and after synchronization, and the blue signal to the z-axis (height) of the reference marker. Before the synchronization, the first MIDI message timestamp (the dashed line at about 8s) is long after the keystroke of the reference key, meaning that the MIDI recording (with Anvil Studio) started before the capture (with Qualisys Track Manager). The solid green line shows the new reference timestamp after the synchronization process, shifted to the onset of the peak (at about 2.5s). The same time shift is applied on the whole MIDI matrix, leading to complete synchronization of motion data and MIDI data, as seen in Section 4.3.4.

![Figure 4.8: MIDI synchronization](image)

### 4.3 Results and Discussion

During the capture sessions, the different setups of markers and cameras were tested with different pianists. Results for the marker models are shown and discussed in Section 4.3.1. Results for the camera setups are presented in Section 4.3.2. Issues encountered with the tested setups, concerning both markers and cameras, are described in Section 4.3.3. Finally, results for the MIDI synchronization are presented in Section 4.3.4.

An example of a capture taken during this project can be found at [http://youtu.be/lNwWH1kJJ4](http://youtu.be/lNwWH1kJJ4). More details about this example can be found in the video description.

#### 4.3.1 Marker Models

No major issue with the markers has been reported by the participants. Despite a slight discomfort just after the markers placement on the body, the pianists became accustomed to them. The bone-based marker model turned out to be more practical than the joint-based model for the performer. Markers placed on phalanges (instead of joints with the joint-based model) are indeed more stable, and can be stuck to the hand with adhesive tape. On the joint-based model, markers must be placed with sticky gum, as tape would lead to a significant inconvenience on the joints, as it would not efficiently embrace their bending movements. According to the pianist’s hands and skin, the sticky gum was more or less effective, and sometimes a marker moved slightly because of repetitive flexions and extensions of limbs, or even fell during a capture session, forcing to a second iteration of the capture.

---

3. The vibrations perceived after the keystroke all along the capture are due to the fact that the bearing of the digital keyboard is not perfectly robust, resulting in a small vibration of the keyboard, and hence the marker placed on it, while playing.
A second practical drawback of the joint-based model for the performer is that markers on fingertips can touch the keys in some rare cases, especially the edges of the black keys, higher than the white ones. Apart from this minor inconvenience, the two models allowed a correct representation of the hand.

Still, in spite of its practical drawbacks, the joint-based marker model was preferred for our specific application, because of its better representation of certain important features. Fig. 4.9 shows hand skeletons re-created with Qualisys Track Manager for both models. The joint-based model (on the left) allows a better representation and visualization, as the created bones (in dark grey) correspond to the real bones. Angles of each joint can easily be extracted from this model. On the bone-based model (on the right), the created bones are somewhat abstract, and the real joints are not represented.

Figure 4.9: On the left: Bone-based marker model skeleton - right hand. On the right: Joint-based marker model skeleton - right hand

As we will see in Chapter 7, a PCA of two renditions with the two different marker models leads to similar results. Although the joint-based model provides a better representation of certain features, the bone-based marker model is thus as efficient as the joint-based model to extract consistent features of the hand, in terms of PCA.

The main problem encountered in the marker setup was the lack of visibility and comfort of the fingertip markers with the joint-based model, but their utility, except for visualization, can be put into question. As a matter of fact, as the DIP-joint movement is greatly linked to the PIP-joint movement, there is not much additional information provided by a marker placed on the tip of the finger.

4.3.2 Camera Setups

The two marker models were tested with the 6-camera setup. As expected, the results of the captures showed that the joint-based marker model was more complicated to capture, and the 6-camera setup was not sufficient for efficient capture with this model. Several cameras have been added and the results showed an improvement in the capture quality. The final camera setup was the 12-camera setup as shown on Fig. 4.6 and in Appendix B.2. This setup allowed efficient capture of every marker of the hands, even in the more extreme cases. Incidentally, the cameras placed behind the pianist also allowed covering the whole body of the pianist, and markers were thus added on the back of elbows and on the neck for visualization, or any possible analysis prospect.

Although the 12-camera setup is more efficient than the 6-cameras one, it is much more complicated to achieve. As a matter of fact, the challenge of a good positioning of every camera is not trivial. The point is that even if no camera is directly in front of each other, reflexions of cameras can be seen on the keyboard. This issue was encountered especially with the following pairs of cameras: #2-5, #3-4, #7-12, #8-11 (see Fig. 4.6). A particular case was camera #6 that is just above the keyboard and could see its own reflection on it. For this particular case, the solution was to slightly move the camera and change its orientation, so that it was not aimed directly at the keyboard. For the other
cameras, the solution was also to slightly change camera orientations, to deviate as much as possible the reflections away from other cameras lenses. As an example, the reflection of camera #5 through the piano aims right between cameras #2 and #7, avoiding major perturbation for those two cameras. Another solution can be to moderately change heights, and distances from the cameras to the piano, to deviate or diminish perturbation for other cameras. Although these complications could be managed for the 12-camera setup, the question must be raised about a further increase of the number of cameras to improve the capture quality. More cameras would indeed allow covering more observation planes (see Fig. 4.4), but would risk to lead to more perturbation for other cameras. Some interesting additional cameras positions could be under cameras #2, #3, #4 and #5, at the height level of cameras #9 and #10. It would allow a better cover of a hemisphere above the piano. These cameras could for instance improve the capture of the thumb. But they would also lead to important perturbation between each other.

4.3.3 Setup Issues

Concerning disappearance of markers in a capture, two main issues were encountered. The first one involves the thumb, and was encountered during renditions of the piece Für Elise (see database description, Chapter 5). The thumb is globally the most difficult finger to capture, because of its different orientation, but also because of the thumb crossover technique, sometimes resulting in the disappearance of the last markers of the thumb under the hand. This issue was generally well managed with both camera-setups, and more efficiently with the 12-camera setup. Nevertheless, in some cases the thumb disappeared a time too long to apply an efficient interpolation process.

The second disappearance issue concerns fingertips, with the joint-based marker model. This issue was encountered for several participants on the piece Comptine d’un autre été (see database description, Chapter 5.). Fig. 4.10 shows a particular position of the right hand. The thumb and pinky are hitting black keys, and the index finger a white key, successively and at a fast rate. As a consequence, the index finger must be highly bent to hit efficiently the white key. This particular position leads to a disappearance of the index fingertip marker, besides the fact that the marker can be bothersome for playing. Even the cameras #9 and #10 (see Fig. 4.6) could not see the marker in this extreme case. Nonetheless, its occultation time, i.e. about a keystroke time, is short enough for a accurate polynomial interpolation.

![Figure 4.10: Right hand particular position during Comptine d’un autre été](image)

4.3.4 MIDI Synchronization

The MIDI synchronization process was verified by visualizing data of a particular exercise used to study pianists’ hand control (see Chapter 5). Fig. 4.11 shows a segment of a motion capture of this exercise. The blue signal corresponds to the height of the right thumb fingertip hitting 16 times the same key at a certain rhythm. The green vertical lines show the MIDI timestamps of the corresponding keypresses. We can see that each MIDI timestamp appears as the thumb is going down, and corresponds to a sudden slowdown of the thumb, indicating the onset of contact with the key. This
verification process was done on many captures and showed good results for all of them, meaning that the synchronization method is efficient.

![Figure 4.11: MIDI synchronization test - exercise extract (right thumb)](image)

The method used for MIDI acquisition in an external program (Anvil Studio), and its posterior synchronization with motion data, is simple and effective. Nevertheless, it is not compatible with real-time performance including gesture recognition and sound synthesis.

4.4 Conclusion

The present chapter presented an important part of this research. It aims to help in future projects by giving a view of this field of motion capture research. It shows that a setup is not an obvious issue, although few research papers dwell on the subject.

This chapter presented different approaches to the motion capture setup issue. It proposed and discussed two different marker models for the hand: the bone-based and the joint-based models. Both models allow a global representation of the hand. Depending on the goal of the capture, one model or the other can be more appropriate. The joint-based model is more representative, but also more cumbersome and complicated to capture. For a real-time acquisition and performance context, the bone-based model or another simpler model thus seems more adequate. For both models, the choice of markers was limited to the available equipment for the research, and prospects for improvement can be envisaged. The 12-camera setup or a similar one could allow the use of 4 mm hemispherical markers on the entire hand, making the pianist freer.\(^4\)

Two camera setups were proposed and discussed in this chapter. The first model was composed of six cameras, and the second one of twelve cameras. More cameras allow a more accurate and precise motion capture, by covering more observation planes, and taking into account more occultations, and extreme cases. Nevertheless, a trade-off must be made on the number of cameras, as too many cameras induce more risks of perturbation for each other. The camera placement is highly dependent to the intended capture and study, and needs to be carefully designed to be effective.

The global setup developed in the context of this project was effective, and the combined usage of the marker model, the cameras configuration, and the MIDI acquisition and synchronization methods allowed the collection of a database of high quality presented in Chapter 5, exploitable for the analysis carried out in Part III. An example of data captured with the final setup can be seen at: [http://youtu.be/lNhMHlkjNJA](http://youtu.be/lNhMHlkjNJA).

---

\(^4\) For the future user, the practical drawback of these markers must be noted: apart from being less visible, they may be less robust than bigger markers, and are not particularly easy to manipulate.
Chapter 5

Presentation of the Database

5.1 Introduction

This chapter focuses on the presentation of the database collected with the setup presented in Chapter 4. The development of the database was another important part of the present project, and resulted from several sessions organized for the collection of the data. For its development, four pianists with different expertise levels have been solicited. These pianists are presented in Section 5.2.

The choice of the exercises or pieces to be played by the participants is very important, as it has a great influence on the relevance of the analysis. For this purpose, a new piano exercise has been specifically designed to answer several questions concerning the hand motor control of pianists'. This exercise is presented in Section 5.3, and its analysis is presented in Chapter 7.

As explained in Chapter 4, a major goal of this project was the capture of actual piano pieces to analyze real and complex pianists’ expert gestures. For the relevance of the analysis, the selection of pertinent pieces is also essential. For this reason, six different piano pieces, from different repertoires and with different peculiarities have been selected. These piano pieces are presented in Section 5.4, and their analysis, based on PCA, is presented in Chapter 7.

All the captures were achieved with the placing of markers on all the upper part of the body including arms, shoulders, and the head. All the captures were recorded with a frame rate of 100 FPS.

5.2 Presentation of the Pianists

Four pianists have been solicited for the acquisition of the data. In the rest of this report, they will be called "pianist 1" to "pianist 4". Table 5.1 briefly presents characteristics of the these pianists.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Sex</th>
<th>Laterality</th>
<th>Age (Years)</th>
<th>Practice Period (Years)</th>
<th>Intensity</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pianist 1</td>
<td>M</td>
<td>Left</td>
<td>22</td>
<td>17</td>
<td>Occasional/Academic(7)/Autodidacte(10)</td>
<td></td>
</tr>
<tr>
<td>Pianist 2</td>
<td>M</td>
<td>Right</td>
<td>17</td>
<td>9</td>
<td>Occasional/Academic</td>
<td></td>
</tr>
<tr>
<td>Pianist 3</td>
<td>F</td>
<td>Right</td>
<td>23</td>
<td>11</td>
<td>Intensive/Academic</td>
<td></td>
</tr>
<tr>
<td>Pianist 4</td>
<td>F</td>
<td>Right</td>
<td>28</td>
<td>21</td>
<td>Intensive/Academic</td>
<td></td>
</tr>
</tbody>
</table>

To study the motor control of pianists, participants with different expertise levels have been recorded. Pianist 1 is an occasional pianist, mainly self-educated, and has 17 years of practice. Pianist 2 has only 9 years of occasional practice and has thus the lowest expertise level. Pianist 3 followed an intensive training for 11 years. Finally, pianist 4 is the only career pianist, and thus supposedly the most skilled, with 21 years of intensive training. Except for pianist 1, each pianist was solicited for
a single recording session. The first two sessions, during which pianists 1 and 2 were recorded, were achieved with the use of the bone-based marker model, as at that time the joint-based model had not yet been tested. For the next two sessions, during which pianists 3 and 4 were recorded, the joint-based marker model was used. A fifth session was also organized, during which pianist 1 was again recorded, but that time with the joint-based marker model on the hands.

5.3 Hand Motor Control Exercise

The exercise developed for the analysis aimed for several goals in the analysis of the pianists’ hand motor control. The first goal was an objective comparison between each finger in a basic keystroke movement, and a comparison between both hands. The second goal was to investigate the effect of the tempo on this movement. The third goal was to see if pianists are able to play this basic movement in different ways, i.e. using different redundant DOF of the arm and hand.

To achieve this analysis, the exercise must allow getting an adequate database, including simple and individual movements of every finger, in sufficient numbers to calculate reliable statistics. In this exercise, the pianist was therefore asked to stroke the same key sixteen times with each finger individually, following the tempo of a metronome. This exercise allowed easy obtaining of a lot of samples for each finger playing, while other fingers were still. It thus allowed getting a general model of the hand, for each finger individually playing. Moreover, the exercise was designed for both hands. In all the previous studies, the analysis was generally limited to the right hand. This exercise allows studying the balance between the motor control of each. As we will see in the analyses, this feature plays an important role in the expertise level of a pianist. Another advantage of this exercise was the easy extraction of a segment corresponding to one finger playing, for individual statistics calculations. This could be done by using the timestamps of the MIDI data recorded together with the movements.

The exercise followed a simple note progression, for each hand, from DO to SOL (C major scale), so that the hand could stay at the same place above the keyboard. The arm and the wrist could hence stay at the same level, while each finger was hitting a key. For more information, a piano score and its description can be found in Appendix C.1.

To observe the influence of the speed of execution on the movement, the participants were asked to play the exercise twice naturally, following a metronome at two different tempi. For investigation on the independent control of redundant DOF of the upper limb, the exercise was repeated with the following restrictive instructions on the movements. The pianist was first asked to play the exercise naturally. Then he was asked to use only flexion-extensions of the fingers for the whole exercise, then only flexion-extensions of the wrist, and finally only flexion-extensions of the elbow. Table 5.2 summarizes the different instructions for the exercise, in the order they were given to the pianists.

<table>
<thead>
<tr>
<th>Tempo (BPM)</th>
<th>Movement instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>180</td>
</tr>
<tr>
<td>2</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
</tr>
<tr>
<td>5</td>
<td>120</td>
</tr>
</tbody>
</table>

The exercise was achieved by several pianists with different levels (see Section 5.2), to extract possible differences in the results due to their expertise level. The analysis of hand motor control is presented in Chapter 6. As MCP joints are used in this analysis, the appropriate marker model to capture this exercise was the joint-based model. As pianist 2 was only captured with the bone-based model, his performance was not taken into account in the analysis.
5.4 Piano pieces

The second part of the database collected in the framework of this master’s thesis consists of extracts of real piano pieces. Several extracts were asked to the four pianists:

- Bagatelle No. 25 in A minor (Für Elise), Ludwig van Beethoven, bars 1-24;
- Prelude and Fugue in C Major, BWV 846, Johann Sebastian Bach, bars 1-11;
- Fantasy in C Major, Op. 17, 1st Movement, Robert Schumann, bars 1-18;
- Comptine d’un autre été, Yann Tiersen, bars 5-16 ("verse"), bars 17-21 ("chorus");
- The Promise, Michael Nyman, bars 3-8 (main repetition);
- Struggle for Pleasure, Wim Mertens, bars 27-30 (main repetition).

In the rest of this report, pieces will be labeled with the composer’s name. Most of these pieces were chosen as they are famous, and hence more likely to be known by pianists. Each one of these pieces has its particularity, and its own complexity for both hands. This set of pieces therefore covers a wide range of music complexity or variety, making comparisons more interesting, and analysis more relevant.

Two main types of pieces can be distinguished in this list. The first three pieces are famous classical pieces for solo piano. The three other pieces are modern piano pieces. Among the three classical pieces, some differences can be noted, in a hands movement point of view: Beethoven uses quite various patterns for right hand, and simple arpeggios for left hand. Bach uses arpeggios for both hands. Schumann uses complex and fast patterns for left hand, and mainly octaves for right hand. Among the six pieces, Schumann can be seen as an exception, as it is the only piece where the left hand game is much more complicated than the right hand for the pianist.

The three selected modern pieces can be classified in the music style "minimal music" [43]. This music has the particularity to be mainly composed of very repetitive patterns. Mertens has also the particularity to use simple and symmetric patterns for both hands.

All these pieces are analyzed in Chapter 7. This analysis is based on the PCA method explained in Section 3.4. A table giving details about the captures of the pieces can be found in Appendix C. It shows which pianist played which piece, and details about the durations and the setup used for the captures.

5.5 Conclusion

The construction of a large database was an important part of this project. The developed database consists of data of four pianists, each one recorded on an exercise, and several real piano pieces. The exercise allowed getting movements of each finger of both hands played a large number of times individually. For each pianist, five occurrences of the exercise were recorded to capture different types of movements, due to different instructions and tempi. Six piano pieces have been selected, and each one has been recorded by one or several pianists, allowing obtention of a wide range of pianists’ expert gestures, for different levels of expertise. All the captures include trajectories for all the upper part of the body, recorded at 100 FPS.

This database has another major asset, as all the gestures have been recorded together with MIDI data. Indeed, these MIDI data, synchronized with the gestures, allow precise and powerful annotation. They can serve as landmark, and they can be used for simple extraction of different musical passages. Moreover, they allow recreation of the sound resulting from the pianists’ gestures. In conclusion, the database developed for this research is easily reusable for further analysis or any other use, thanks to the MIDI data.
Part III

Analysis
Chapter 6

Hand Motor Control

6.1 Introduction

In the present research, two different analyses have been conducted. The first one is a classical analysis of the pianists' hand motor control, based on the exercise described in Chapter 5. This analysis is inspired from many previous papers on the subject, for which various exercises were designed to answer specific questions regarding pianists' hand motor control (see Chapter 1). The second one is more innovative, as real complex piano pieces have been analyzed, thanks to PCA. Both analyses have their utility, and are somehow complementary. The first one is quite basic and is based on statistical moments directly calculated on the captured data. The advantage of this type of study is that results are easy to interpret and focused on an exercise developed to answer specific questions. PCA is on the other hand more difficult to interpret, based on a transformation of the data, but it is more general and powerful. It indeed allows extraction of several relevant features from any complex piano piece. These features can be used for analysis, but also allow dimensionality reduction in HCI.

In the present chapter, we focus on the exercise-based analysis. The second analysis is presented in Chapter 7. The exercise designed for this study (see Chapter 5) aims to answer several questions. First, we investigate how pianists naturally play a basic keystroke with each finger, without specific instructions. A comparison is made between the control of each finger individually. To observe the influence of the tempo on their natural playing technique, they have been asked to play at two different tempi. They also have been asked to play the exercise following different instructions about their movements. The purpose of these instructions is to observe how pianists can play the same pattern by using different redundant DOF of the upper limb.

Our exercise has another peculiarity, as it has been designed for both hands (see piano score in Appendix C.1). Most of the previous studies were only focused on the right hand of the participants. Yet, in music, both hands may often have different roles in the playing of the instrument. An investigation of the balance between both hands motor control can therefore be interesting. In piano playing, both hands supposedly have the same role, i.e. executing keystrokes. It is thus very different from instruments like the guitar, where one hand is used to press chords, and the other to pluck strings. However, on the piano, the role of each hand is actually very different. The main difference between both hands is due to the pitch. The piano has the particularity to have an extremely wide range of frequencies, allowing playing at the same time very low and very acute sounds. Yet, the low pitches are located on the left of the piano and are generally played with the left hand. Respectively, the acute sounds are played with the right hand. In an harmony, acute pitches are generally used to play the melody (e.g. flutes, violins), while low pitches are used for accompaniment (e.g. a bass). On the piano, the left hand thus has somehow the role of a bass, and the right hand the role of a flute. Their roles may have a great impact on the way of playing, and therefore on their training for complex gestures. For a career pianist, both hands must be highly trained, and a good balance is imperative to master any complex piece and give expression with both hands. For a less trained pianist, the left hand may be neglected, as its accompaniment role is much simpler than that of the right hand.
Both exercise-based and pieces-based analyses will try to confirm that hypothesis. In the exercise-based analysis presented in this chapter, it is done by developing an exercise for both hands, and comparing statistical moments separately calculated for each hand. The method is presented in Section 6.2, results are presented in Section 6.3, and Section 6.4 presents a discussion of the analysis.

6.2 Methods

The exercise has been designed to analyze movements of the upper limb for each finger individually playing keystrokes. In this exercise, each finger is played sixteen times individually while others are still, so as to have a lot of samples of a basic keystroke movement with the same finger, to extract reliable statistical values corresponding to each finger individually. The repeated playing of the same finger allows avoiding the influence of the movements of transition between fingers that occur all the time in a classical exercise like a note progression.

All data were processed and analyzed with MATLAB. Several scripts and functions were written for segmentation and analysis. The database includes all occurrences of the exercise performed by three pianists (pianist 1, 3 and 4, see Appendix C). Fig. 6.1 shows the block diagram of the hand motor control analysis. The input data are matrices containing marker trajectories, and matrices containing the MIDI messages recorded from the keyboard during the capture. There is one trajectory matrix and one MIDI matrix for each capture, i.e. for each occurrence of the exercise with a particular instruction. For each occurrence of the exercise, the pair of matrices (trajectories and MIDI messages) is previously synchronized as explained in Chapter 4. The analysis consists in two steps. First, each trajectory matrix is segmented thanks to the corresponding MIDI matrix. Each segment corresponds to the capture period while a finger is individually playing. Statistics are then calculated from each segment, and features are extracted and compared.

![Figure 6.1: Hand Motor Control Analysis](image)

6.2.1 Matrix segmentation

Before the analysis process, the trajectory matrix is segmented in ten smaller matrices, each one corresponding to the marker trajectories when one finger is individually played. One segment of data is created for each finger, between the first MIDI timestamp and the last one corresponding to this finger. As each finger is played on a single key, it is easy to find the corresponding MIDI messages by looking at the MIDI pitch (note number). To avoid taking into account transition movements between fingers, the data just before the first timestamp of a finger, and just after the last one are not extracted. These two periods can be considered as transition periods.

6.2.2 Feature Extraction

For each segment, several features are extracted from the trajectories. To analyze the principal redundant DOF playing a role in the basic keystroke movement, angles of the elbow, the wrist and MCP joints are extracted. All joint angles are extracted by calculating angles between lines connecting
the markers placed on the corresponding adjacent bones. For each segment, the Standard Deviations (SD) of the joint angles are calculated. Each SD is thus based on sixteen samples of the same keystroke movement. These values indicate how much each joint is solicited, and therefore give a profile of the global movement of the arm in the keystroke movement. The z-axis coordinates of fingertips markers are also extracted, from which means and SD are calculated, to give a profile of positions and movements of the hand and fingers.

6.3 Results

6.3.1 Natural Playing

Joints Solicitations

Fig. 6.2 shows the joints solicitations for an occurrence of the exercise. This is the case of pianists 1, 3 and 4 playing the exercise at a tempo of 120 BPM (Beat Per Minute, meaning Inter-Onset Intervals (IOI) = 250 ms), with a natural playing, and with the right hand. For each segment, angles standard deviations (ASD) were calculated for the MCP joint of the playing finger (blue), the wrist (green), and the elbow (red). All results can be found in Appendix D.

Figure 6.2: Angles standard deviations of redundant degrees of freedom - Pianists 1, 3 and 4, 120 BPM, natural playing, right hand

We can see that the three pianists have quite different profiles for the natural playing at 120 BPM. Pianist 4, the more trained pianist, uses more elbow and less MCP joint than the two others to execute the exercise. On the opposite, pianist 1 uses mainly his MCP joints. Pianist 3 is intermediate between pianists 1 and 4. Different pianists can thus naturally use very different techniques in piano playing, by solicitation of different redundant DOF of the arm and hand.

Fingertip Heights

Fig. 6.3 shows statistics for fingertip heights for the three pianists executing the exercise with a natural playing at a tempo of 120 BPM, with the right hand. Standard Deviations (SD) (graphs on

---

1. For instance, the elbow angle is obtained by calculating the angle between the lines corresponding to the upper arm and the forearm. The upper arm line is formed by two markers placed on the shoulder and the elbow, and the forearm line by two markers placed on the elbow and the wrist.
the left) are an image of the movements of each finger during a segment, i.e. when one finger is playing individually. For pianist 4, the SD of the five fingers are decreasing from the thumb to the pinkie, regardless of the playing finger. This is due to the fact that pianist 4 rested her pinkie on the keyboard while playing with the rest of the fingers. For pianists 1 and 3, we can observe differences according to the finger playing. It is not necessarily the playing finger that has the highest SD. A look at a slow-motion replay of the exercise performed by these pianists allowed observing that at the bottom of the stroke of a key, the playing finger was suddenly blocked and thus stopped its descent, while adjacent fingers continued their inertial descent.

Means of the fingertip heights (graphs on the right) show general poses of the hand during a segment. We can easily observe that the playing finger is always the lower than the other fingers. By comparing the three pianists, we see that pianist 4 generally has a lower pose than that of the others, as her hands rested on the keyboard during the exercise. Pianist 1 used another technique: each not-playing finger seems to be as high as possible, especially the pinkie. The thumb is generally lower, because of its particular biomechanics.

**Tempo Influence**

The natural playing at a tempo of 180 BPM (IOI = 167 ms) shows major differences compared to the 120 BPM tempo for pianists 1 and 4, but not for pianist 3 (see Appendix D). At this tempo, the arm is generally stiff because of strong and rapid muscular pulses. For pianist 1, ASD are greater for the wrists, and for pianist 4, ASD are greater for the MCP joints. Again, different strategies in the use of redundant DOF are employed for the same purpose.

Fingertip heights statistics for 180 BPM show different results especially for the playing finger.
The means show that the playing finger is generally lower than at 120 BPM, and the SD show that the playing finger is also more still, meaning it does not raise at all at 180 BPM. This is not the case for not-playing fingers. The amplitude of the playing finger is probably reduced to manage the very short IOI at that tempo. Different findings were noted by Dalla Bella and Palmer (2011) [21]. They found that at a faster tempo, the finger height was generally higher, aiming a better temporal accuracy. However, those results concerned a very different exercise, i.e. a simple melody. In our exercise, the same finger makes a keystroke movement on each beat of the tempo, and hence does not have time to rise between keystrokes. It is not the case in a melodic exercise where each finger makes keystroke movements alternately with the other fingers. The real tempo for an individual finger is therefore much lower in these exercises. Moreover, the temporal accuracy issue in a melodic exercise involves transitions between fingers. It is not the case in our specific exercise.

Hand Influence

For each pianist, results of the left hand are generally very similar to those of the right hand. The three pianists seem to have a high symmetry in natural playing between both hands at a tempo of 120 BPM. At a faster tempo, light differences between both hands appear for the three pianists.

6.3.2 Motor Control of Redundant Degrees of Freedom

![Figure 6.4: Angles standard deviations of redundant DOF - Pianist 4, 120 BPM, right hand - above: fingers playing, middle: wrists playing, below: elbows playing.](image)

In addition to the natural playing, each pianist was asked to perform the exercise while restraining his movements. Fig. 6.4 shows results for pianist 4 playing the exercise trying to move only the fingers, the wrist, and the elbow respectively, as instructed. As expected, major differences are observed for the three instructions. The playing with the finger (first graph) is highly visible for all fingers, except for the thumb. We can indeed see that the ASD of each MCP is high (about 10°) while that of the wrist and elbow are very low. The case of the thumb can be explained by the fact that it is a particular finger, as explained in Chapter 2. The playing with the wrist (second graph) and with the elbow (third graph) are also visible, as compared to the first graph ASD of the MCP are lower and that of the wrist and elbow respectively increase with the corresponding instruction. The instructions were thus well followed by pianist 4, meaning that she can easily play the same pattern with different redundant DOF. The same observations for pianist 1 and 3 (see Appendix D) show that both of them also followed the instructions quite properly for both hands. In sum, despite the abstract nature of the instructions,
each pianist understood and followed them properly and could easily use different redundant DOF for the same keystroke movement.

Hand Influence

Results for the left hand are generally similar to those for the right hand. Nevertheless, a global comparison of both hands for the three pianists, tend to show that the most trained pianist (pianist 4) has a better symmetry between both hands with the three instructions, and hence has a left hand motor control as fine as her right hand motor control. This corresponds with the hypothesis that to handle a wider range of complex piano pieces, a more skilled pianist must have a good balance of training of both hands.

6.4 Discussion

This exercise, and the extracted features aimed to answer different questions. The results shown in Section 6.3 allow answering some of them.

Without instructions, the three pianists seemed to execute the exercise with very different approaches. One of them used his fingers more and another one used more his elbows. These results correspond with several previous studies. Dalla Bella and Palmer (2011) [21] showed that pianists have their own personal signature concerning their finger kinematics in piano playing. This personal signature can surely be extended to the whole upper limb, as it would also correspond with findings made by Lourenço (2007) [44] about different "piano interpretation schools" using different dynamics in piano playing. According to this paper, pianist 1 would play more in a French interpretation tendency, using essentially his hands and fingers, and pianist 4 in a Russian interpretation tendency, making use of the "weight" technique of the arms. However, the database used for this analysis, i.e. three pianists including only one career pianist, limits the possibilities of interpretation of the results.

The three pianists seem to have a good motor control of redundant DOF. They can play separately with each DOF, meaning there is a fine decoupling of these DOF. Pianists can therefore freely play with different techniques and are not tied to a single personal way of playing the piano. A comparison with non-skilled pianists would be interesting in order to observe how much the training plays an influential role in this decoupling, and hence how training helps dealing with the enslaving effects of the hand biomechanics (see Chapter 2). This finding presents an interest for researches in the field of new human-instrument interfaces, led for instance by the CIRMMT and Numediart.

Results are available for both hands. In general, pianists have a good balance between both hands motor control. Nevertheless, the career pianist (pianist 4) showed a slightly finer balance, as at a level of concert pianist, the left hand is as important as the right hand. In Chapter 7, we will see that pianist 4 used an especially high number of principal movements with the left hand on the piece of Schumann, particularly hard for that hand. This result also indicates a fine decoupling in her left hand DOF, and hence corresponds with her particularly good motor control of the left hand.

The exercise was especially developed to be able to extract reliable statistics for each finger, thanks to a high number of successive keystrokes with the same finger. SD and means were extracted for each finger from each segment. The means allowed observing the global pose of the hand while playing keystrokes with each finger. SD on the other hand, did not lead to any important conclusion. Further investigation could be done by extracting more relevant features for finger comparison. Correlations between fingers were calculated, but did not lead to any conclusion. As a matter of fact, SD or correlations between finger movements depend on many parameters like inertia of not-playing fingers, different finger lengths, different enslaving effects, making comparison difficult.

As for any of the previous studies on this subject, these results are highly dependent to the exercise developed for the analysis. Different exercises have been used in other studies, such as successive
octaves\textsuperscript{2} \cite{3}, tremolo\textsuperscript{3} \cite{20}, short note progressions\textsuperscript{4} \cite{19}, or a simple melody \cite{22}.

For all these exercises, several instructions must be given, for instance by indications on a musical score. The use of these indications allows the study to be well scoped, as we know exactly which movements are analyzed. One of the instructions generally given is the tempo. Another important instruction generally provided concerns the melodic style, and is given by the terms legato or staccato. Staccato means that the successive sounds must be short and detached from each other. On the opposite, legato means that the successive sounds must be linked, resulting in a more harmonious sound. These terms have an incidence on the way of playing: to play staccato, i.e. detached notes, a finger making a keystroke must be taken off the key before the next finger begins another keystroke. On the opposite, to play legato, another finger can begin a keystroke before the previous one is finished. In our specific exercise, the indication was the staccato, as it allows raising the hand from the keyboard between each keystroke, while legato forces to have all the time a finger on the keyboard. Staccato is therefore more adapted to study the control of redundant DOF.

The problem of this additional indication is that the analysis becomes even more specific to the exercise. All studies involving a particular exercise are therefore limited to it, and results can hardly be generalized. For our specific exercise, particular additional instructions were given to the pianists, to test their control of redundant DOF. A problem of this type of instructions is that they are open to interpretation if they are not precise enough. Results may thus be dependent to the interpretation made by each pianist. It is therefore important to be very precise when giving the instruction to the performer.

6.5 Conclusion

In this analysis, pianists’ hand motor control was investigated. We found that naturally, pianists can play a same exercise in completely different ways. We observed that the three pianists who participated in this study have a good separate control of redundant DOF of the upper limb, allowing them to easily play the same pattern differently. The way of playing a pattern has an impact on the resulted sound. However, depending on the instrument, this impact is not so relevant. This skill of musicians to vary their ways of playing a pattern can be exploited in the design of DMI to amplify their influence to the synthesized sound. This is done by creating extended instruments, and the use of such instruments is called augmented music \cite{1}.

The proposed study allowed answering some questions about the natural playing and control of redundant DOF of pianists. However, we only obtained preliminary results, and prospects of a further analysis based on the same exercise can be envisaged with an extended database, including more career pianists, and some novice pianists. A sufficient number of participants could allow investigation of the influence of expertise on hand motor control. We could also observe statistically significant differences between different groups of pianists, allowing extraction of tendencies in personal natural playing.

This type of analysis, based on a specific exercise as done in several previous studies, has some limitations. The results are generally restricted to the specific exercise used for the analysis, and are often not extendable to general movements. Chapter 7 proposes a totally different approach that allows analysis of real complex piano pieces. This analysis based on PCA, is not limited to one single exercise, and is therefore more general.

\textsuperscript{2} An octave is the interval between two sounds whose one sound’s fundamental frequency is the double of the other’s. E.g., the simultaneous sound of a Do3 (261.63 Hz) and a Do4 (523.25 Hz) is an octave. To perform an octave on piano, the thumb and pinkie make a simultaneous keystroke (e.g. the thumb on Do3 and the pinkie on Do4).

\textsuperscript{3} The tremolo technique results from fast alternate keystrokes of two notes, with two different fingers (e.g. alternate keystrokes of the thumb and the pinkie \cite{20}).

\textsuperscript{4} Note progression: successive keystrokes of each fingers in an ascending order.
Chapter 7

Principal Component Analysis

7.1 Introduction

This chapter focuses on the second analysis of pianists’ expert gestures, which is based on PCA. In order to have a reliable and meaningful database for this analysis, several pianists were asked to execute extracts of several piano pieces (see Chapter 5). These performances were captured with the motion capture setup described in Chapter 4. In order to completely capture the complex movements of the pianists’ hands, about 25 markers had to be placed on each hand. The captured data thus comprised about 75 components (three per marker), to represent the movement of one hand, the latter actually consisting of 23 DOF. All the components hence provided much redundant information about the hand movement. Moreover, most of these 23 DOF are not independent, due to several enslaving effects described in Chapter 2. A question can therefore be raised about the complexity of pianists’ expert gestures, and the number of components needed to correctly represent them. PCA can be used to precisely answer that question.

That issue presents several different interests. First, understanding the operation of the body in complex movement of course presents a great interest in biomechanics and biodynamics research fields. This information can indeed be used for many applications such as medicine, ergonomics, sport, music and all disciplines requiring complex movements of the body. The second interest of this analysis is to investigate how much the dimensionality of data representing complex movements of the hand can be reduced, in order to facilitate HCI. In HCI, a recent tendency is the control of a computer by the use of movements of the body, and particularly the hands. However, these movements are complex, and it is not straightforward to virtually model them, nor to transmit to the computer the relevant information they provide. To simplify this process, PCA offers two solutions. First, it allows the reduction of data representation to fewer (principal) components, for information transmission with as little data as possible. Additionally, each PC alone provides some relevant information, meaningful for the modeling of the hand. Each PC indeed represents an orthogonal movement of the hand, while before the PCA process, a component only represents the insignificant movement of one point on one axis. In this chapter, we will see how complex movements such as pianists’ expert gestures can be represented with a reduced dimensionality. The number of components needed to represent one movement depends on its complexity, and hence can depend on different parameters. On the one hand, a complex music piece can require more complex or varied gestures. On the other hand, a more skilled pianist can execute more complex movements to execute any music with more precision and personality. The impact of the complexity or variety of musics, and that of the skill of a pianist, on the gestures complexity will also be investigated in this chapter.

The balance of motor control of both hands can also influence the complexity of gestures performed by each hand. This balance has already slightly been investigated in Chapter 6, but this analysis will provide more precision on that issue. The methods used to perform the extraction of the data, and the process of PCA, are described in Section 7.2. Results are presented in Section 7.3, and the analysis is then discussed in Section 7.4.
7.2 Methods

Fig. 7.1 shows the block diagram of the analysis. First, The MIDI data, previously synchronized with the marker trajectories (see Chapter 4), are used as landmarks for extraction of reliable data from the captured pieces. From these pieces extracts, the marker trajectories of each hand are extracted, to be separately analyzed. The PCA process is then applied on each hand, and PC are calculated. The results are seen in the form of eigenvalue ratios, telling how much information is provided by each component, and in the form of videos, showing the projections of the movement on the first few PC.

![Figure 7.1: Principal Component Analysis](image)

7.2.1 Reliable Data Extraction

Before the analysis, reliable motion data must be extracted from all the captured data. The motion data are reliable if each important marker was correctly captured by the motion capture system, if the pianist played without any mistake during the capture, and if they are comparable with data extracted from other captures. A correction of the marker trajectories, based on polynomial interpolation, can be done in Qualisys Track Manager to fill small holes in trajectories, before exportation into MATLAB. In MATLAB, the extraction of correct and comparable data is done by the use of the corresponding MIDI data, previously synchronized with motion data, as landmarks. The MIDI data are indeed easy to decipher, and allow a comparison with a piano score to verify if the extract was correctly played, and a comparison with other captures, to extract identical musical passages (see Fig. 7.2).

In Fig. 7.2, we represent extracts of Tiersen played by three pianists, with the corresponding MIDI matrices. Each point corresponds to a played note, i.e. a line in the MIDI matrix. It is represented with its onset timestamp on x-axis and its MIDI pitch on y-axis. Observation of the three graphs and comparison of MIDI pitches with the piano score allows verifying if each pianist correctly played the extract. Comparison between the three graphs allows identifying identical musical passages. The extraction of comparable motion data is done by keeping the data between two timestamps corresponding to the same sequence of played notes for each pianist. For example, the extraction of the musical passage corresponding to the chorus of Tiersen (bars 17-21 on the piano score, see Chapter 5) is shown in Fig. 7.2. The lengths of the extracted data for the three pianists will thus respectively be: 17.25 s, 17.23 s and 20.06 s. Lengths and others details about all the other extracts that were analyzed in the present research can be found in Appendix C.

7.2.2 Hands Data Extraction

The trajectory matrix includes all the markers placed on the body during the capture. In order to separately analyze the movements of each hand, the trajectories of the markers placed on the hands
must be extracted from the matrix. Each hand will thus be associated with a trajectory matrix with less dimensions. During this analysis, we focused on hand movements only. Hence, we moved the marker coordinate system, initially placed on the piano, on the central marker of the wrist. The movements are hence not relative to the keyboard anymore, but relative to the wrist.

### 7.2.3 Principal Component Analysis

The PCA algorithm used in this analysis is a modified version of the algorithm provided by the MOCAP Toolbox. The latter only allowed an analysis based on the covariance matrix between components. The correlation matrix has been used instead, to process normalized data and hence give the same weight to each component during the analysis.

The input matrix, i.e. the trajectory matrix, contains the coordinates of each marker of the hand all along the capture. For the joint-based model, the number of markers analyzed for one hand is 26. The number of components is therefore $3 \times 26 = 78$. For the bone-based model, the number of components is $23 \times 3 = 69$. The correlation matrix is calculated between these components. The eigenvectors and eigenvalues are extracted from the correlation matrix, and the data are represented with new components (in the new orthonormal base) defined by the eigenvectors (see Section 3.4). Projections on this orthonormal base are also calculated to visualize which movement is represented by each new component. In order to correctly visualize all PC, a new MATLAB function was written to create videos that allow simultaneous visualization of the principal movements, also called eigengestures (as coined in [43]), corresponding to each of the first PC. Examples of these videos can be found in the results Section 7.3.
7.3 Results

Fig. 7.3 shows the PC of the movement of the right hand of pianist 1 on the chorus of Tiersen, obtained from the trajectories of the markers of the joint-based model. The curves shown in Fig. 7.3 are difficult to interpret, but each one represents a tangible movement, as explained below. These movements can be visualized on this video: http://youtu.be/6qT1hn9UUJ4. The weight of each component, i.e. the amount of information it provides, is shown in Fig. 7.4.

![Figure 7.3: Principal components - Pianist 1 - Right hand - Tiersen chorus (100 FPS). The first PC is the one that shows the largest variability.](image)

![Figure 7.4: Engeivalue ratio of the first principal components - Pianist 1 - Right hand- Tiersen chorus](image)

In Fig. 7.4, we have arbitrarily kept only the principal components that provide at least 95% information about the movement. We can observe that in this case the first eight components are sufficient to represent the hand movements to 95%. In other words, the other 70 components provide only very little additional information, and are therefore of little use.

The red curve shows the information accumulated by the first components. It can be seen that this curve has a high gradient for the first three components, and then flattens out. The first three components indeed represent 77.6% of the movement. The next component only provides about 6% additional information.

Fig. 7.5 shows a screen-shot of the video. The original movement of the hand is shown in the upper left corner, in green. In order to visualize the movements properly, the hand is shown with two different perspectives. The first eight PC are displayed in order, from left to right and from top to bottom. Traces of each marker allow seeing the movement which each component represents. For instance, the first component approximately represents the inclination of the hand. The fingers and the palm are indeed going up and down relative to the wrist, throughout the extract. These up-down movements are also respectively coupled with slight right-left movements. We recommend the reader

1. The trace of the first PC is represented in green.
to rely on the video (see http://youtu.be/6qIHn9UWc5w) for a better visualization.

![Hand Principal Component Analysis](image)

**Figure 7.5:** Original and 8 projections on the first principal components - Pianist 1 - Right hand - Tiersen chorus

For all other captures, cumulative information provided by the first PC can be found in Appendix E, in the form of Cumulative Eigenvalue Ratio (CER) curves. A general observation of these graphs shows that eight PC are generally sufficient to represent movements of the hand with 90-95% of accuracy.

Table 7.1: Numbers of PC required to provide 95% of the information for: Left hand - Right hand.

<table>
<thead>
<tr>
<th>Joint-based model</th>
<th>Bone-based model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pianist 1</td>
<td>Pianist 3</td>
</tr>
<tr>
<td>Tiersen (Chorus)</td>
<td>7 - 8</td>
</tr>
<tr>
<td>Tiersen (Verse)</td>
<td>7 - 8</td>
</tr>
<tr>
<td>Beethoven</td>
<td>5 - 8</td>
</tr>
<tr>
<td>Bach</td>
<td>6 - 8</td>
</tr>
<tr>
<td>Shumann</td>
<td></td>
</tr>
<tr>
<td>Nyman</td>
<td></td>
</tr>
<tr>
<td>Mertens</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1 gives, for each capture, the number of PC needed to represent the movement with at least 95% information, for each hand. We can see that depending on the pianist, the piece, and the hand, this number varies between 5 and 10.

### 7.3.1 Models Comparison

To compare the models (joint and bone-based hand models), we must rely on pianist 1, as it is the only pianist who tested both models. The comparison can be based on Table 7.1, or on the CER graphs (see Appendix E). A look at this table, in the columns of pianist 1, for both marker models, shows that the results are quite similar. For the joint-based model, a few more PC are needed to catch
95% of the movement. Fig. 7.6 shows the CER curves for the four pieces extracts played by pianist 1 with both models (Tiersen verse and chorus, Beethoven, and Bach), for the right hand. The results are quite similar, but slight differences can be noticed. Except for Tiersen verse, the curve is generally higher with the bone-based model, meaning the first PC catch a higher ratio of the information. Nevertheless, it must be noted that 100% of the information contained in one model, is not necessary the same information as that contained in another model. For instance, the total information provided by a model composed of two markers is significantly lower than that provided by a model with plenty of markers, allowing the capture of each DOF. Furthermore, it must also be noted that differences in the curves can be due to the fact that the pianist did not exactly play the same way during both captures.

A comparison of the videos showing the projections on the PC with both models can provide more precision. The video showed earlier (http://youtu.be/6qIIfn9UUJw4) was created from a capture with the joint-based model. For the bone-based model, a video can be found at http://youtu.be/yRj_n5AdtLs. We can see in these videos a good correspondence between both models for the movements represented by each component. The information ratio of PC 3 and 4 seems to be switched, but the rest is quite identical. The comparison for the left hand provided good results too: each component was corresponding except for PC 5 and 6 whose information ratios were switched.

### 7.3.2 Hands Comparison

The CER graphs shown in Appendix E allow a direct comparison of both hands for each capture. Depending on the piece, the movement of the left hand is generally supposed to be simpler than that of the right hand. For example, in the extract of Beethoven chosen for the analysis, the left hand only produces simple arpeggios with three fingers, while the right hand produces a varied melody with five
fingers. Results for the three pianists show that less PC are needed to represent the movement of the left hand, than that of the right hand. It is shown by a generally higher curve.

Schumann has been chosen for its particularity to require a highly skilled left hand, in contrast with Beethoven. We can see that for the capture of the pianist 4 playing Schumann, the CER curve of the left hand is lower than that of the right hand, and 10 PC are needed to catch 95% of the movement. The analysis therefore seems to confirm the complexity of the left hand on this piece. Another piece that revealed to be more complex for the left hand is the chorus of Tiersen, as we can see on the CER curves corresponding to the four pianists. This extract is indeed highly simple for the right hand, while the left hand requires quite complex movements, as explained in Section 7.3.4.

Mertens, performed by pianist 1, is particularly interesting for hands comparison, as identical patterns are used with both hands to play this piece. Both CER curves are particularly high as the gestures are simple and repetitive. Nevertheless, the curve of the left hand is higher for the first PC. It may be an indication of a higher coupling between the DOF of the left hand for pianist 1.

7.3.3 Pianists Comparison

The comparison of pianists 1, 3 and 4 can be based on four pieces extracts, played with the joint-based model. The CER curves seem to be generally lower for pianist 4. It must be noted that pianist 4 is a career pianist, more trained than the two others. The CER curves of pianist 4 seem to show her finer motor control in piano playing, allowing her to play with more precision and give more expression in her performance.

The balance between both hands seems to be connected with the level of training of the pianists. For pianists 1 and 3, the CER curve for the left hand is generally higher, except for Tiersen chorus where the movements of the right hand are simple, using repetitive patterns with only three fingers. For pianist 4, the CER curves of the left hand are quite low, compared to other pianists. Pianist 4 thus seems to have a finer balance between the motor control of her hands. This statement corresponds with the analysis of hand motor control carried out in Chapter 6, and with the hypothesis that a professional pianist must have a good balance between both hands to master complex pieces, and give more expression with both hands.

Pieces Concatenation

The concatenation of several extracts\(^2\) allows more general results, not dependent to a single piece. Fig. 7.7 shows the cumulative information provided by the first eight PC for both hands of the pianists 1, 3 and 4, on the concatenation of four extracts: Tiersen chorus and verse, Beethoven, and Bach. For all the CER curves, 8 to 10 PC are needed to provide 95% information, except for the left hand of pianist 1, whose CER curve is particularly high. In this case, the first PC already contains 60% information on the movement, and 6 PC are sufficient to represent the movement to 95%. This curve probably indicates a lack of training of the left hand compared to the right hand. It must be noted that, unlike pianists 3 and 4 that followed an academic intensive training, pianist 1 is mainly self-educated. Moreover, the repertoire played by these pianists is quite different: pianists 3 and 4 have followed a classical education, while pianist 1 has played in a more modern repertoire like jazz and variety, where the right hand is generally complex but the left hand is often based on simple patterns like chords or octaves.

7.3.4 Pieces Comparison

Most of the selected pieces are very different in several points of view, and a comparison based on PCA can thus be complicated. Moreover, the results for a same piece may differ with the selected

---

\(^2\) The concatenation of several extracts is done in MATLAB by concatenating trajectory matrices. The resulting matrix is then used in the whole process.
Figure 7.7: Cumulative Eigenvalue Ratios - Concatenation of four extracts: Tiersen chorus and verse, Beethoven, and Bach.

extract and its length, making the comparison delicate. Nevertheless, several observations can be noted.

It can be seen that, in general, a varied melody leads to a lower CER curve, meaning that more PC are required to represent the movement. For example, the chorus of Tiersen is simpler than the verse on the right hand. The patterns of the chorus are indeed repetitive arpeggios played with only three fingers. Similar simple patterns are used on the left hand of Beethoven, Nyman, and the two hands of Mertens. We can notice that the minimalist musics we selected for analysis, especially Mertens, have quite high curves, meaning they use simple patterns, according to PCA.

However, among these minimalist pieces, we can notice that Tiersen is not so simple. The pattern of the left hand, although repetitive, requires rather complex movements formed by sequences of harmonic intervals\(^3\), requiring a fine dexterity. Moreover, for the right hand, although the chorus is very simple and typical of minimalist music, the verse is more varied, resulting in a lower CER curve.

Schumann is, as expected, a complex piece according to PCA, as it leads to low CER curves, especially for the left hand.

7.3.5 Principal Component Projections

According to the patterns used for a particular piece, movements may be very different, and hence the first PC may be very different for several pieces. Videos of the projections on the eight first PC were therefore created for each capture, and allowed seeing which movements were found in the most representative components.

Observation of these videos showed that the first PC often represented similar movements for different pieces. PC 3 to 5 generally represented global movements of the hand, involving coupled up-down, left-right, opening-closing movements of the hand. The pronation-supination of the forearm\(^4\) were often represented by PC 4, 5 or 6 according to the piece and the pianist. PC 6 to 8 generally represented movements involving one, two or three fingers, more dependent to the particular piece.

---

3. An harmonic interval is the simultaneous keypress of two notes.
4. The pronation-supination of the forearm induces rotation of the hand around the axis going through the forearm, and is therefore considered as a movement of the hand itself. This movement was often coupled with alternate flexion-extension of the thumb and the pinkie, as the purpose of both movements is similar, that is alternate keystrokes of extreme fingers.
7.4 Discussion

7.4.1 Methods and algorithm

The methods and the algorithm of the analysis worked well for all the captures, but several choices can be discussed, and improved.

Four pianists have been captured for the analysis. A higher number of pianists would allow getting more precision on the statements set in the analysis. Nevertheless, the few pianists selected presented different skill levels, hence allowing interesting comparisons and observations. However, pianist 2, who was the less trained pianist, did not have a sufficient knowledge of the selected extracts, and all the captures were therefore not kept for analysis.

The sample of different played and analyzed pieces was on the other hand quite large. Three pianists performed four different extracts, from three different pieces: Tiersen, Beethoven and Bach. These pieces were mainly selected as they are well-known by most pianists. Pianist 1 also played two minimalist pieces: Mertens and Nyman. Pianist 4 played an extract of a master piece in the classical repertoire: Schumann. All these pieces allowed getting a wide and eclectic set of pianists’ expert gestures. The concatenation of different pieces allowed very general comparison of the pianists.

All the MATLAB functions written for the analysis were effective. The creation of the video was particularly interesting, as it allowed the simultaneous visualization of the original movement and the contribution of each PC to the movement. This organization of the visualization allowed easy comparison of each PC in a single piece, but also to compare the eigengestures obtained for different pieces, different pianists or different models.

Several choices were made for the execution of the algorithm. Each hand was separately extracted, from the wrist to the fingertips. For each hand, the coordinate system was placed on the wrist, to analyze the precise and complex movements of the hand itself. The movements of the rest of the body were thus not taken into account in the analysis. An analysis of the movements of the arms or of the whole body with PCA could be an interesting future work.

In the PCA process, the correlation matrix was used rather than the covariance matrix. All the data were thus normalized for the analysis. The reasoning for this choice was to give the same weight to each component of the movement, as in music a small movement can have a great impact on the resulting sound and is therefore as important as a large one. Nonetheless, the algorithm was tested with the use of the covariance matrix, and showed quite similar results. The CER curves were a bit higher, and the projections gave similar movements.

7.4.2 Utility of the analysis

This analysis allowed providing many general results regarding pianists’ expert gestures. Several statements concerning the use of PCA in the study of gestures can be set:

- first, it allowed great reduction of the number of dimensions required to represent the movement of the hand, even for highly complex movements like piano playing. 3 or 4 PC indeed generally represent 80% of the movement, and 8 PC were sufficient to represent 90-95% of the movement;
- secondly, the eigenvalue ratios seemed to provide a profile of variety, or complexity of the movement executed during a capture. This profile could be exploited for analysis and comparison of different pianists, or pieces. However, the interpretation of the results must be taken cautiously, as they can depend on several factors, like the length of the analyzed extract, the chosen passage, the comfort of the pianist, and capture conditions. Moreover, the definitions and differences between variety or complexity of a movement themselves can be open to interpretation;
finally, the projections of the movement on the PC allowed decomposition of the movement and visualize its most relevant parts. This movement decomposition could also be exploited in a future work for further analysis and characterization of gestures.

PCA is therefore a very powerful tool for several purposes in motion analysis. In the context of the present project, it allowed comparison of different pianists and characterize their expertise level. It also allowed comparison of different piano pieces played for the analysis, and characterization of their complexity or variety, for both hands.

We noticed an important parameter concerning the expertise level of a pianist: the balance of motor control between both hands. This information about the balance of a pianist can be useful in education. For example, we could diagnose a bad balance for pianist 1 as his control of the left hand seemed to be limited. An adequate compensatory training could be the practice of the piece Schumann, whose left hand’s CER curve is particularly low, meaning this piece is more skill-demanding for the left hand.

Additionally, a comparison of the joint-based and the bone-based model showed that they exhibit similar performance regarding the representation of the hand. The CER curves were indeed similar, and more importantly the projections on the PC gave the same movements.

7.5 Conclusion

In this chapter, we could analyze pianists’ movements in real piano pieces. PCA allowed decomposition of the pianist’s expert gestures in eigengestures. We saw that less than ten eigengestures are sufficient to correctly represent the hand gestures of the pianists in real performance. It means that the 23 DOF of the hand are not completely independently used even in the practice of highly complex gestures by highly skilled players. This statement can be connected to Chapter 2, explaining different causes of limitations of the movements. PCA can show how much movements are limited, and could therefore be an indication on the condition of the enslaving effects on a pianist’s movements. In particular, the neuroplasticity could be observed through PCA, as this is the main enslaving effect that can be dealt with by training.

The methods used in the analysis, based on PCA, could easily be extendable on any other movement. Analysis of other musicians could allow comparison of the different eigengestures used in the playing of different instruments. In music, or sport, or any other field where complex movements are used, eigengesture decomposition could allow deciphering the movement, and its observation could allow understanding how to make a perfect movement. For example, it could help a golfer to improve his swing, or prevent a tennisman from injuring his elbow. In conclusion, this analysis method is thus very powerful, and its modularity allows its use in many fields.

PCA is not only useful for analysis, but the process of reduction of data dimensionality can be, and already is, widely used in movement recognition. In the context of this project, we showed that the pianist’s expert gestures can be represented with a few PC. Four components indeed generally represent 80% of the movement, and less than ten were sufficient to represent it to 95%. In the context of the design of DMI, these components could be used to virtually recreate a model of the hand, and use this model as a controller for the synthesis of new sounds.
Conclusion

Conclusion of the Thesis

Several goals were targeted in this research. The first goal was to achieve an efficient setup for the capture of the expert gestures of pianists on entire complex piano pieces. For this purpose, several setups have been singled out and tested. The issue of the marker model has been investigated, and two models have been considered: the joint-based model, and the bone-based model. Tests showed that the joint-based model is more representative, but more cumbersome, and more complicated to capture than the bone-based model. PCA showed that the bone-based model allowed correct extraction of the eigengestures of the hand and is therefore as efficient as the joint-based model for the modeling of the hand. The camera setup has also been investigated. A logical approach has been followed to manually optimize the number and positions of cameras for the specific capture. Two camera setups have been described in this report. The 12-camera setup was the most effective, and allowed the capture of most of the extreme patterns played by the pianist and to take into account most possible positions of the hands during the capture. The 6-camera setup was a bit less effective than the 12-camera setup, but was easier to configure and allowed adequate capture with a smaller number of cameras.

In parallel with motion data, the MIDI data were also recorded. They were indeed very useful in the extraction step of the two analyses performed in the present research. As MIDI data were recorded separately from motion data, they had to be synchronized afterwards. The processing of the MIDI data and their synchronization with the motion data were successfully achieved in MATLAB for each capture.

The second main purpose of the present research was to analyze pianists’ expert gestures. Four pianists were therefore recorded with the motion capture setup specifically developed for this study. The four pianists had different expertise levels, allowing investigation of the impact of the pianists’ skills on their gestures. The pianists were asked to perform an exercise, and to play excerpts of several piano pieces. An important contribution of the present work is hence the collection of a multi-pianist MIDI-annotated motion capture database of several piano pieces and exercises.

The exercise-based analysis aimed at answering several questions concerning pianists’ hand motor control. First, it allowed investigation of how pianists make use of their redundant DOF for the natural playing of a basic keystroke. It also allowed investigation of the independent control of each of the main redundant DOF used in a basic keystroke. Moreover, the influence of the tempo was taken into account in the analysis. Additionally, the exercise was separately executed with both hands and all fingers to examine differences between hands and fingers, regarding their motor control. Results showed that the pianists captured for the analysis could naturally play in completely different ways, and had a quite good independent control of redundant DOF of the upper limb, meaning they can play the same pattern with countless different ways. It may be possible to exploit these skills for the control of new DMI. The design of new extended instruments could indeed enhance the impact of the impressive possibility of movements on the produced music, and thus amplify the faculty of personal expression of musicians.

The PCA-based analysis was much more innovative than the first one, as it allowed study of complete movements performed by pianists on actual complex piano pieces. This analysis therefore had
a great advantage over those restricted on a specific exercise. It allowed us to decompose the pianists’ expert gestures into eigengestures, showing the relevant parts of a movement in piano playing. Eight of these eigengestures were generally sufficient to model 95% of the movements of the hand. Indeed, for most of the pieces captured with several pianists, eight PC generally provided at least 95% information on the movement. Pianists’ gestures are thus not as complex as they could be if the 23 DOF of the hand were all completely independently used. However, the analysis seems to indicate that more trained pianists can better take profit of the possibilities offered by the hand DOF. Indeed, they use more different eigengestures in piano playing, probably meaning that their neuronal zones are better adapted to the use of more different movements, allowing more precision and hence more personal expression on piano.

Future Work

Further Development of the Project

Several future improvements of this project could be considered. First, the database was limited to four pianists, and only one career pianist. A broader database could allow obtaining more significant results from both analyses. This broader database would also encourage a deeper statistical analysis based on the exercise. We could extract tendencies in piano playing, determine how much pianists play differently according to their expertise level, their piano school, their age or their laterality (does a left-handed pianist play differently from a right-handed pianist?). The power of PCA could also be more deeply investigated by the analysis of a broader database, including more pianists playing identical pieces.

Nevertheless, the collected database is already significant, and allows prospects of further analysis of pianists’ expert gestures. The analysis could be extended to the rest of the body that was captured, i.e. the arms, shoulders and the head. Different analyses could also be conducted on the data, using for example different algorithms for dimensionality reduction like Fisher’s linear discriminant analysis [46]. Gesture segmentation and recognition processes using for instance hidden Markov models [47] could as well be tested on it.

Applications to Human-Computer Interaction

From the results of our analyses, several prospects of future work can be imagined in the context of HCI. The exercise-based analysis shows the extent of the possibility of making different movements, which could be used for amplified interaction with a computer in the context of music (DMI), or for the control of any other interface. The PCA-based analysis shows that the movements of the hand, which is the main body part used for interactions with machines, can easily be represented via PCA, even for highly expert gestures like piano playing. The PC can be used to virtually model hands’ movements, and from this model control different machines, like extended instruments.

The design of new DMI could be based on the use of eigengestures of an existing instrument. The mapping of each eigengesture used to control the instrument could directly be linked with one parameter controlling the sound synthesis. It would allow an orthogonal and thus independent control of each parameter, and thus a more efficient control of the sound.

Applications to Biomechanics and Medicine

The PCA-based analysis conducted in the present research could easily be applied to different types of musicians, or to athletes. It could also be used for other complex movements, performed in the industry. It is thus possible for each discipline to extract and visualize the used eigengestures. Accordingly, one could classify those disciplines depending on their complexity. One could also categorize those disciplines according to specific performed eigengestures.
Observation and comparison of eigengestures for several people in a sport could provide a better understanding of movements, and bring to light the reasons why some athletes are better than others in their discipline. For example, the comparison between the swing of a professional golfer and those of learners could help the latter make progress, by correcting the wrongly executed eigengestures. Most frequent wrong movements could be determined for all sportive and musical disciplines, or any other discipline requiring complex movements. In parallel, analysis of athletes injured by strenuous bad movements could allow identification of the wrong movements causing their injury. More generally, it would help physiotherapists knowing more about how people get injured because of wrong movements.

Furthermore, we showed that PCA can provide a profile of the complexity of movements, and is thus correlated with neuroplasticity. In a clinical context, these profiles could thus be used for assessment of the neural condition of patients suffering from cerebral palsy or any other motor disability. Particularly, it could be a powerful tool for monitoring the evolution of those diseases. PCA could also help to identify the eigengestures that are most neglected by the patient, and accordingly imagine adapted rehabilitation exercises.
Bibliography


Appendix A

PCA: An Illustrative Example

A basic example of principal components extraction is presented here, allowing a visualization of the process, and thus a better understanding.

Fig. A.1 shows a random 2D data set generated with MATLAB. It is represented with two components, on axes $x_1$ and $x_2$.

Before performing the process, the data set is centered on its center of gravity, see Fig. A.2.

The covariance matrix is then processed, and the eigenvalues and eigenvectors are extracted.

The resulting eigenvalues ($l$) for this random data set were: $\lambda_1 = 37.27, \lambda_2 = 2.73$. The weighted amount of information provided by each component $i$ is given by $\lambda_i/(\sum_{i=1}^{d} \lambda_i)$. The first component provides thus $\frac{37.27}{40} = 93\%$ of information and the second provides only $7\%$.

The two eigenvectors $v_1$ and $v_2$ are represented in Fig. A.3, respectively in green and red, and are the new axes
of representation of data. The data set can be represented by keeping only the first component, i.e. on the first
eigenvector (see black dots, Fig. A.3):

```matlab
Xpc1=Xc*v(:,2); % (in MATLAB, the biggest eigenvalue is the last)
```

This representation allows a reduction of dimensions (from two to one), losing only 7% of the information of
the data set. The component, or axis $v_1$ is the axis of representation leading to the minimum difference between
the blue dots (two dimensions representation) and the black dots (only one dimension representation on $v_1$), in
the least square sense.

This process can be generalized for more than two components. For a hand (having 23 DOF) represented with
26 markers, the number of components is $3 \cdot 26 = 78$ (x,y,z-axis for each marker). The Chapter 7 shows that the
components can be reduced to less than ten, while keeping 95% information, even for a complex piano piece.

Figure A.3: Axis Transformation
Appendix B

Motion Capture Setup

B.1 6-Camera Setup

Fig. B.1 shows a photo of the 6-camera setup, corresponding to Fig. 4.5.

Fig. B.2 shows the markers captured by each camera of the 6-camera setup (see Fig. 4.5). The index number of each camera is written in the lower-left corner of each view, and the number of marker seen by each camera is written in the lower-right corner.

Fig. B.3 shows a zoom on the sixth camera’s view.

Fig. B.4 shows the thumb under technique, during a rendition of Für Elise by pianist 1, using the bone-based markers model, and captured with the 6-camera setup.

B.2 12-Camera Setup

Fig. B.5 shows a photo of the 12-camera setup, corresponding to Fig. 4.6.

Fig. B.6 shows the video views of each camera of the 12-camera setup (see Fig. 4.6). The index number of each camera is written in the lower-left corner of each view.

Figure B.1: 6-camera setup
Figure B.2: 6-camera setup - cameras views corresponding to the setup Fig. 4.5.

Figure B.3: 6-camera setup - zoom on camera 6

Figure B.4: Thumb under technique on the left hand, during Für Elise rendition of pianist 1 - seen by the 6-camera setup - bone-based model.
Figure B.5: 12-camera setup

Figure B.6: 12-camera setup - cameras views corresponding to the setup Fig. 4.6.
Appendix C

Database Details

C.1 Hand Motor Control Exercise

Fig. C.1 shows the exercise developed to analyze hand motor control of pianists. The exercise can be described as follows:

First, on the left hand, the pinkie plays sixteen times a DO₂, then the ring finger plays sixteen times a RE₂, and so on until the left thumb plays sixteen times a SOL₂. Then the right hand plays, from DO₃ to SOL₃, respectively with the right thumb to the right pinkie. This exercise is repeated with different instructions on the tempo (120 or 180 BPM\(^1\)), and on movement restrictions (play naturally, or only by using a certain limb).

Figure C.1: Hand motor control exercise. Each finger is played 16 times on the same note (DO to SOL for each hand), at a tempo of 120 or 180 BPM, and respecting some additional movement instructions.

\(^1\) A tempo of 120 BPM corresponds to IOI of 250 ms, and a tempo of 180 BPM corresponds to IOI of 167 ms.
C.2 Pieces Extracts

The pieces asked to the pianists were adapted to their expertise levels. For instance, pianist 2 only played Beethoven and Tiersen (and only Tiersen was kept for analysis). On the opposite, Schumann was asked to pianist 4 only, as it requires a high expertise level. Table C.1 summarizes the pieces extracts analyzed in the present research. These passages were extracted from long captures, during which the pianist played in loop, to ensure a part with a high capture quality, and without any playing error. For a same piece, identical musical passages were extracted for every pianist, to keep comparable data.

Table C.1: Analyzed captures characteristics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Capture</th>
<th>Date</th>
<th>Time (s)</th>
<th>Markers Model</th>
<th>Cameras Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>Beethoven</td>
<td>24/04/14</td>
<td>7.14</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Tiersen (Verse)</td>
<td>24/04/14</td>
<td>15.24</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Tiersen (Chorus)</td>
<td>24/04/14</td>
<td>17.12</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Bach</td>
<td>24/04/14</td>
<td>39.03</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Nyman</td>
<td>24/04/14</td>
<td>15.07</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Mertens</td>
<td>24/04/14</td>
<td>18.04</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p1</td>
<td>Beethoven</td>
<td>08/05/14</td>
<td>6.14</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p1</td>
<td>Tiersen (Verse)</td>
<td>08/05/14</td>
<td>16.10</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p1</td>
<td>Tiersen (Chorus)</td>
<td>08/05/14</td>
<td>17.25</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p1</td>
<td>Bach</td>
<td>08/05/14</td>
<td>34.27</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p2</td>
<td>Tiersen (Verse)</td>
<td>28/04/14</td>
<td>17.34</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p2</td>
<td>Tiersen (Chorus)</td>
<td>28/04/14</td>
<td>17.02</td>
<td>Bone-based</td>
<td>6</td>
</tr>
<tr>
<td>p3</td>
<td>Beethoven</td>
<td>09/05/14</td>
<td>7.12</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p3</td>
<td>Tiersen (Verse)</td>
<td>09/05/14</td>
<td>17.07</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p3</td>
<td>Tiersen (Chorus)</td>
<td>09/05/14</td>
<td>17.23</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p3</td>
<td>Bach</td>
<td>09/05/14</td>
<td>39.16</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p4</td>
<td>Beethoven</td>
<td>12/05/14</td>
<td>8.00</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p4</td>
<td>Tiersen (Verse)</td>
<td>12/05/14</td>
<td>19.19</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p4</td>
<td>Tiersen (Chorus)</td>
<td>12/05/14</td>
<td>20.06</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p4</td>
<td>Bach</td>
<td>12/05/14</td>
<td>39.21</td>
<td>Joint-based</td>
<td>12</td>
</tr>
<tr>
<td>p4</td>
<td>Schumann</td>
<td>12/05/14</td>
<td>30.20</td>
<td>Joint-based</td>
<td>12</td>
</tr>
</tbody>
</table>
Appendix D

Hand Motor Control - Results

D.1 Principal Redundant Degrees of Freedom

The graphs below show ASD of the MCP joint playing (blue), the wrist (green), and the elbow (red) for the three pianists that participated to the exercise captures (pianist 1, 3 and 4). Each pianist executed the exercise five times, as explained in Appendix C. Note: In the graphs below, the ASD for the wrist (green) and elbow (red) have been multiplied by 2, for better visualization and comparison.

Pianist 1
### D.2 Fingertips height

The graphs below show SD and means of all fingertips (thumb to pinkie from blue to red), for each segment (i.e. while each fingertip is individually playing).
Pianist 1 - 120 BPM - natural - Left Hand

Pianist 1 - 120 BPM - natural - Right Hand

Pianist 1 - 180 BPM - natural - Left Hand

Pianist 1 - 180 BPM - natural - Right Hand

Pianist 3 - 120 BPM - natural - Left Hand

Pianist 3 - 120 BPM - natural - Right Hand

Pianist 3 - 120 BPM - natural - Left Hand

Pianist 3 - 120 BPM - natural - Right Hand

XI
Pianist 3 - 180 BPM - natural - Left Hand

Pianist 3 - 180 BPM - natural - Right Hand

Pianist 3 - 180 BPM - natural - Left Hand

Pianist 3 - 180 BPM - natural - Right Hand

Pianist 4 - 120 BPM - natural - Left Hand

Pianist 4 - 120 BPM - natural - Right Hand

Pianist 4 - 120 BPM - natural - Left Hand

Pianist 4 - 120 BPM - natural - Right Hand

Pianist 4 - 180 BPM - natural - Left Hand

Pianist 4 - 180 BPM - natural - Right Hand

Pianist 4 - 180 BPM - natural - Left Hand

Pianist 4 - 180 BPM - natural - Right Hand

XII
Appendix E

Principal Component Analysis - Results

The graphs below show the CER curves of the first eight PC for each capture. A low slope of the curve indicates more complex or various movements, as more components are then needed to represent them.

Joint-based Model

![Graphs showing CER curves for each pianist and composer]
Bone-based Model

Pianist 1 - Nyman - Hands PCA

Pianist 1 - Mertens - Hands PCA
Principal Components

Pianist 1 - Tiersen chorus - Hands PCA

Pianist 1 - Tiersen verse - Hands PCA

Pianist 1 - Beethoven - Hands PCA

Pianist 1 - Bach - Hands PCA

Pianist 2 - Tiersen chorus - Hands PCA

Pianist 2 - Tiersen verse - Hands PCA

XV