POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

Master's Degree Thesis

A serious-game-based pipeline for rehabilitation and assessment of upper limb impairment

Supervisors

Prof. Gabriella OLMO

Prof. Gianluca AMPRIMO

Dr. Claudia FERRARIS

Candidate

Daniele PILI

July 2025

Summary

The progressive ageing of the population has resulted in an increase of neurodegenerative disorders, including Parkinson's Disease. In order to ensure the preservation and the enhancement of motor functions, continuous rehabilitation programs would be necessary. However, these programmes are often costly and not conceived for home-based use.

Exergames, videogames designed to perform therapeutic physical exercises, represent a cost-effective solution suitable for telerehabilitation. They enable continuous patient monitoring and increase engagement throughout rehabilitation process.

This study analyzes the data collected by the *Palestra* exergame, developed by CNR-IEIIT, which employs a single RGB camera combined to Google MediaPipe Pose for human pose estimation. The system is designed for parkinsonian subjects, and focused on upper limb motor recovery, through simple arm-raising exercises (single arm, alternating and simultaneous movements, performed both laterally and frontally).

The analysis was performed using a dedicated automatic pipeline, designed to extract angolar trajectories and a set of parameters aimed to evaluate the user's motor performance. In order to assess the quality of the acquired data and to define limits and potential of the exergame, a validation procedure was carried out. This validation was done against an optoelecronic system and included a total of five healthy subjects. Moreover, the system was tested on parkinsonian patients with different degrees of motor impairments.

The results show that the pipeline produces reliable angular trajectories in both healthy and parkinsonian users, thus allowing the extraction of robust parameters in both cases. The Bland-Altman analysis conducted on the entire trajectories demonstrates a good agreement in lateral raises. Temporal parameters, such as the Peak-to-Peak Time (PPT), exhibit a high concordance between the two systems (Bias= 0.00 s, LoA=$[-0.09\,\text{s}, 0.09\,\text{s}]$). Angular parameters, like Range of Motion (ROM), present marginally lower agreement (Bias= 2.03 °, LoA=$[-9.16\,°, 13.22\,°]$). In frontal tasks, the Bland-Altman analysis reveals slightly wider discrepancies in temporal parameters, even though they remain comparable with those identified

in lateral tasks (PPT: Bias= 0.01 s, LoA=$[-0.23\,\text{s}, 0.24\,\text{s}]$). Larger differences are observed in the angular parameters (ROM: Bias= $-3.56\,°$, LoA=$[-23.53\,°, 16.42\,°]$). Nevertherless, these discrepancies, both temporal and angular, are still considered accetable for the intended context of use.

In conclusion, *Palestra* represent a promising solution for telerehabilitation, due to its ease of use, cost-effectiveness and ability to collect data in a continuous and non-invasive manner. The accuracy levels of the system, as demonstrated by the validation analysis, are accetable and suitable for the intendend telerehabilitation applications.

Acknowledgements

I would like to express my heartfelt thanks to Dr. Claudia Ferraris and Prof. Gianluca Amprimo for their irreplaceable guidance during this work. Without their help, I would not have been able to complete my academic path, and for this I am truly grateful.

I am also grateful to Prof. Gabriella Olmo for the opportunity to carry out this work.

I wish to express my gratitude to the Consiglio Nazionale delle Ricerche, for allowing me to use the laboratory, which was essential for the development of my thesis.

I would like to thank my mother, my father and my sister, for supporting me even from miles away.

I am truly thankful to Federica, my love, for always encouraging and helping me to carry on with my studies.

Table of Contents

List of Tables

List of Figures

Acronyms

PD
Parkinson's disease

QoL
Quality of Life

SG
Serious Game

ER
Exergame Rehabilitation

VR
Virtual Reality

HPE
Human Pose Estimation

GMP
Google MediaPipe Pose

SMD
Standardized Mean Difference

CI
Confidence Intervals

RCTs
Randomized Clinical Trials

NMA

Network Meta-analysis

TUG

Timed Up and Go test

UPDRS

Unified Parkinson's Disease Rating Scale

FOG

Freeze of Gait

MoCap

Motion Capture

CNN

Convolutional Neural Network

GUI

Graphical user interface

HA

High Amplitude

LA

Low Amplitude

ROM

Range of Motion

PPT

Peak-to-Peak Time

OR

Over-Reaching Amplitude

PPM

Peaks Per Minute

RV

Raising Velocity

PF

Peak Frequency

ΔHA

High Amplitude Difference

ΔLA

Low Amplitude Difference

ΔROM

ROM Difference

IPID

Inter Peak Interval Difference

PTL

Peak Time Lag

CCL

Cross Correlation Lag

ΔRV

Raising Velocity Difference

ΔMV

Mean Velocity Difference

Chapter 1

Introduction

1.1 General context: the role of exergames in telemedicine

Over the past few years, the progressive aging of the population has entailed a significant increase in neurodegenerative diseases and associated motor disabilities, such as stroke and Parkinson's disease (PD). The incidence of these conditions is particularly high in low and middle-income countries, where access to traditional treatments is often limited [1]. To ensure, as far as possible, an improvement in and a sustained quality of life (QoL) for patients, a continuous rehabilitation process would be recommended. However, conventional rehabilitation often involves high costs, making it inacessible to many. In this context, the use of exergames represents a sustainable solution, as it enables rehabilitation, or at least motor function training, to continue directly at home, thereby reducing costs and providing an effective rehabilitation strategy suitable even in economically disadvantaged areas [2].

Although different authors use the term "exergame" with different connotations, it typically refers to a video game that requires the physical movement of the player to work, thereby promoting motor learning and physical activity [3, 4]. Exergames can be classified as a subgenre of *serious games* (SG), which are games primarily designed for training or educational purposes, with entertainment being a secondary feature [5, 6]. Due to this characteristics, exergames have a wide application in the field of motor rehabilitation.

Exergaming rehabilitation (ER) encompasses a wide spectrum of rehabilitation methodologies that utilize interactive digital environments to support the functional recovery of the patient. ER can be implemented using different devices, including consoles or computers with non-immersive video games, semi-immersive hybrid systems, and fully immersive virtual reality (VR) systems [7].

ER finds a wide application in upper limb rehabilitation. Its primary goal is to enhance upper limb mobility through game-based tasks, aimed at improving or preserving specific motor functions, such as range of motion, coordination, and motor control. These exercises, characterized by an engaging and interactive approach, can be flexibly adjusted according to the patient's motor condition, enabling an effective and personalized adaptations of task difficulty. This approach allows the patient to play an active role in the rehabilitation process, significantly increasing motivation for therapeutic treatment [2, 8].

Most of the proposed exergames represent an effective telemonitoring system. Through the integration of 3D body tracking systems, it is possible to collect quantitative data in a simple and non-invasive manner, providing information on the patient's mobility (such as the trajectories of the identified joints) and their performance in completing rehabilitation tasks (such as a score obtained after completing a task). This continuous and non-invasive data acquisition enables optimal planning of rehabilitation sessions, facilitating targeted and personalized interventions tailored to the patient's specific needs, both motor-related and organizational or personal.[2]

In recent years, exergames have also been used in the prevention and treatment of pathological conditions such as obesity and cardiovascular diseases, as well as to improve postural control and balance in the elderly population. They have also shown potential in rehabilitation of cognitive function, both in the presence of acquired and congenital deficits, such as memory, attention, verbal fluency, learning, problem solving and planning. Indeed, due to their immersive and interactive nature, exergames promote motor learning processes and stimulate experience-based neural plasticity through high-intensity repetition of task-based exercises, progressively increasing task difficulty, real-time feedback, motivation, rewards and even transfer effects. This latter aspect indicates the ability to transfer the improvements achieved in a specific rehabilitation task to different functional domains that were not directly trained.[9]. For example, a patient who regularly plays an exergame based on a tennis game improves not only his motor skills and reaction speed required by the game itself, but also his ability to manage motor and cognitive activities in daily life that are not directly trained, such as memory or decision-making skills.

1.2 Current Technological Limitations and Objectives

Despite the growing interest in the use of exergames for motor rehabilitation in telemedicine, several limitations still hinder their widespread adoption in clinical practice. Firstly, there is considerable inconsistency in the technological approaches

adopted, particularly concerning the hardware employed. Most studies employ Microsoft Kinect technology, which includes an integrated motion capture algorithm proven to be accurate in various validation studies [10, 11, 12]. However, this technology is no longer manufactured by Microsoft, effectively preventing its large-scale adoption and limiting the development of new exergames based on this system. In order to overcome these technological limitations, current research is increasingly oriented toward the simplification of hardware, favoring configurations based exclusively on RGB cameras. This shift has been enabled by the emergence of advanced Human Pose Estimation (HPE) frameworks such as OpenPose or Google MediaPipe Pose (GMP), which have demonstrated the ability to estimate the human body motion using exclusively RGB images. As a result, these systems are more cost-effective, accessible, and easier to implement. Secondly, a general and standardized pipeline for processing exergame-generated data is still lacking, and no clear set of standard clinical parameters for evaluation has been defined. The definition of a unified and standardized procedure, tailored to the tracking system, exercise type, and target pathology, would enable more extensive and in-depth studies on the clinical validity of exergames.

This thesis aims to contribute to addressing the current gaps in the use of exergames for motor function training and rehabilitation in individuals with Parkinson's disease by focusing on the following main objectives:

- To make use of the exergame developed by CNR-IEIIT, referred to as *Palestra*, as the basis for motion data acquisition and subsequent clinical analysis.
- To define an automated pipeline for the analysis of joint data acquired through exergame and the extraction of clinical parameters for motor assessment.
- To validate the trajectories and parameters extracted by the exergame, which uses Google MediaPipe Pose, by comparing them with those obtained using an optoelectronic system, universally considered the gold standard for motor performance evaluation.
- To evaluate the applicability of *Palestra* on subjects with Parkinson's disease for extracting reliable angular trajectories and outcome parameters.

To achieve these goals, the exergame used in this project implements Google MediaPipe Pose as its motion capture system. MediaPipe Pose is an open-source framework developed by Google, which enables automatic real-time detection of human body landmarks, thus generating a three-dimensional skeletal model [13]. This tool runs with both mobile and desktop devices, making it potentially compatible with any hardware environments. It also promotes standardization in pose detection techniques across exergames, while significantly reducing development costs.

1.3 Thesis Structure

This thesis is organized as follows:

- **Chapter 2** provides an overview of the state of the art concerning the use of exergames in rehabilitation. It briefly presents some types of exergames developed for different pathologies in addition to Parkinson's disease. Moreover, a brief description regarding the markerless motion tracking in rehabilitation is presented. A specific focus will be placed on the Google MediaPipe framework, which serves as the underlying technology of the developed exergame.
- **Chapter 3** describes in detail the materials and methods adopted. It covers the technical setup of the exergame, the development of the automatic data processing pipeline, the validation procedure through comparison with the optoelectronic system, the experimental protocol used for data acquisition from individuals with Parkinson's disease and the outcome parameters extracted.
- **Chapter 4** presents the results obtained from the validation analysis. In particular, it describes the findings regarding the overall 3D trajectories analysis, the mean parameters, and single-repetition parameters, further examining the influence of the segmentation algorithm used. It also reports the results obtained by the analysis performed on parkinsonian subjects.
- **Chapter 5** offers an in-depth discussion of the findings, critically analyzing the strengths and limitations of the developed solution. The clinical implications and future prospects for telerehabilitation are discussed, with an emphasis on problems that emerged during the validation analysis and on potential methodological improvements for future research.
- **Chapter 6** provides a critical summary of the goals achieved and those only partially met. The practical implications of the results in both clinical and telemedicine contexts are explored.

Chapter 2

Background

2.1 Effectiveness of Exergames in Motor Rehabilitation: Literature Review

In this section, the application of exergames in the field of telemedicine has been explored, with particular emphasis on their role in telerehabilitation. The objective is to highlight how exergames represent effective and economical tools capable of covering multiple aspects of motor rehabilitation, not only focused on upper limb interventions. The literature review was conducted by analyzing scientific publications released between 2015 and 2025, mainly obtained from PubMed and Google Scholar databases.

In recent years, numerous studies have highlighted the effectiveness of exergames in motor rehabilitation, particularly in the elderly population and in parkinsonian subjects. Improvements in balance, functional mobility, gait, motor recovery of the upper limbs, and, mode generally QoL have been reported as benefits by following rehabilitation protocols involving exergames.

Harris et al.[14] were among the first to conduct a systematic review and meta-analysis investigating the use of exergames as a rehabilitation tool to improve balance and postural control in individuals with Parkinson's disease and healthy older adults. The review included 11 studies, encompassing a total of 325 healthy older participants and 56 subjects diagnosed with Parkinson's disease. To evaluate the effect between the experimental and control groups, the authors used standardized mean difference (SMD) and confidence intervals (CI). The results indicated a significant improvement in static balance (SMD = 1.069, 95% CI: 0.563-1.576), postural control (SMD = 0.826, 95% CI: 0.481-1.170) and dynamic balance (SMD = -0.808, 95% CI: -1.192 to -0.424) among healthy older adults. Additionally, two specific studies reported improvements in static balance (SMD = 0.124, 95% CI -0.581 to 0.828) and postural control (SMD = 2.576, 95% CI: -1.534-3.599) in

participants with Parkinson's disease.

The systematic review and meta-analysis conducted by Cièslik et al. [15] exmined 52 randomized clinical trials (RCTs), including 3081 participants aged over 60 years, demonstrating that rehabilitation protocols based on exergames with motion capture are statistically significant in improving functional mobility compared to control groups (SMD = -0.70, $p<0.01$). Network meta-analysis (NMA) showed that all outcomes analyzed using motion capture-exergames are comparable to traditional rehabilitation exercises, with particularly evident improvements in the Timed Up and Go test (TUG) compared to standard procedures. The authors attribute these positive results to the immediate and continuous feedback that exergames offer, allowing patients to visualize their movements in real time. The study also states that exergames represent a safe rehabilitation modality, suitable for home-based applications and characterized by an absence of significant side effects.

A systematic review by Garcia-Agundez et al.[16] analyzed 64 publications selected according to specific inclusion criteria. The clinical studies analyzed showed improvements in motor and cognitive performance in patients with Parkinson's disease, comparable or superior to control groups. The authors emphasize, however, the need for further research to connect the parameters extracted from exergames to traditional clinical scales, such as the Unified Parkinson's Disease Rating Scale (UPDRS).

Nuic et al.[17] developed an exergame called *Toap Run*, based on the Kinect system, in which patients are required to avoid obstacles or collect objects using limb and pelvic movements, guided by visual and auditory cues. After 18 sessions, the participants (n = 18) reported high levels of satisfaction, feasibility, and acceptability regarding the exergame. Improvements were observed in clinical parameters such as freeze of gait (FOG), the gait-and-balance scale, and the axial score, with reductions of 39%, 38%, and 41%, respectively. Kinematic gait parameters also showed enhancements, particularly in step length, walking speed, and reduction of double-stance time. However, the study presents methodological limitations, including the small sample size and the absence of a control group.

Pachoulakis et al.[18] introduced a Kinect-based platform, consisting of two exergames (Ballon Goon and Slope Creep game), aimed at Parkinson's patients with mild or moderate symptoms (Hoehn and Yahr stage 1 to 3), without severe postural instability. Both games use auditory and visual feedback, and a score-based reward system. However, clinical data confirming the efficacy of this platform.

Amprimo et al.[19] developed an exergame based on the Azure Kinect platform designed to stimulate both upper and lower limb movements, while simultaneously acquiring quantitative kinematic parameters such as joint angles, center of mass position, velocity, and temporal metrics. Tested on 20 parkinsonin patients and compared with 15 healthy subjects, the study confirmed significant differences in

kinematic parameters, validating the system as an effective tool for complementary assessments and suitable for telemedicine.

Given their proven effectiveness in telerehabilitation, exergames are also used in complex and advanced systems, such as the REHOME platform developed by Ferraris et al [20]. This solution is dedicated to the rehabilitation of subjects affected by neurodegenerative disorders, including Parkinson's disease. The platform allows rehabilitation interventions on the motor, cognitive and and sleep-related functions, supported by a continuous monitoring system of the activities performed by patients. The studies conducted have shown that the system is characterized by a high level of usability, making it suitable also for elderly users. The results of the questionnaires administered to the participants indicate that the gamification component, typical of exergames, represents a key element in maintaining a high level of adherence to the proposed rehabilitation program.

Additionally, in the context of post-stroke rehabilitation, the Kinect-based *Jintronix* exergame has proven to be safe, well-tolerated, and effective in improving movement quality compared to control groups [21].

In conclusion, as highlighted by the studies discussed above, exergames show strong potential from a clinical perspective, demonstrating compatibility with telerehabilitation and personalized medicine. Thanks to integrated markerless tracking systems, exergames allow for the continuous acquisition of quantitative data, enabling consistent patient monitoring. However, despite these encouraging outcomes, certain methodological (e.g., small sample sizes, absence of control groups) and technological (e.g., limited accuracy in motion capture) limitations remain. Addressing these challenges is essential to improving the reliability and clinical validity of exergame-based rehabilitation.

2.2 Markerless motion tracking system in Rehabilitation

A key requirement for the functionality of an exergame is the capability of detecting the user's movements in real time and translating them into inputs interpretable by the game [22]. To achieve this goal, Human Pose Estimation (HPE) techniques are employed, which are defined as "the ability of a system to identify the position of the joint centers of the human body in three-dimensional space or in an image" [23].

Currently, the most accurate technologies for acquiring this type of information are marker-based motion capture (MoCap) systems, still considered the gold standard in terms of precision. In these systems, specific reflective or luminous markers are applied to certain anatomical points of the human body, allowing the construction of a three-dimensional skeletal model. The movement of these

markers is then recorded by a series of infrared cameras strategically positioned in space. Despite the high accuracy guaranteed by these systems, they require a considerable number of cameras to minimize the risk of occlusions of the markers during movement. Furthermore, they require highly qualified personnel both for the correct use of the equipment and for the accurate positioning of the markers themselves on the patient. These factors, combined with the high cost of the required equipment, make marker-based systems expensive, less accessible, and therefore poorly suited for widespread adoption in the field of telerehabilitation. For these reasons, their use is not feasible outside of specialized laboratories and controlled environments [24, 25, 26].

To address these limitations, markerless motion capture systems have found wide diffusion in the exergaming and telerehabilitation fields, which estimate the patient's kinematics using simple RGB or RGB-D (depth) cameras in combination with HPE algorithms. More specifically, RGB-D cameras, such as the Microsoft Kinect cameras, record, in addition to the classic RGB video stream, also the distance of each pixel from the camera itself, thus allowing to estimate the depth. HPE algorithms are generally classified based on the representation adopted for the human body model [27]:

- **Planar**: a representation that uses simple geometric shapes, typically rectangles, to approximate the human body.
- **Kinematic (3D)**: models the positions and orientations of joints and limbs in three-dimensional space.
- **Keypoint (2D)**: a two-dimensional projection of the three-dimensional model, often mistakenly confused with the kinematic representation.
- **Volumetric**: a three-dimensional representation using mesh structures.

Among these, the 3D kinematic and 2D keypoint representations are the most used in exergames and markerless systems. The three-dimensional kinematic representation can also be obtained from a single RGB camera thanks to deep learning algorithms such as BlazePose and MediaPipe, which estimate the depth coordinate (z). More accurate systems utilize RGB-D cameras, which provide enhanced depth estimation, or on triangulation-based multi-camera RGB systems [28, 27].

The keypoint representation requires only an RGB camera and deep learning algorithms, such as OpenPose, to estimate two-dimensional landmarks. These systems are capable of providing real-time estimations when supported by sufficient computational resources (e.g., multi-core CPUs with GPU acceleration) and are well suited for rehabilitation applications that involve gross motor movements [29, 27].

In addition to their use as input systems for exergames, motion capture technologies are widely employed in clinical practice, as they allow biomechanical analysis of movement, generating useful information for planning on targeted therapeutic interventions [30]. In particular, biomechanical analyses are extremely useful in the context of rehabilitation, as they allow continuous and systematic collection of patient data. This method allows for constant and effective monitoring of the patient's progresses or changes, without the need to frequently visit specialized medical centers. This approach also offers significant economic advantages for the healthcare system, as it makes it possible to more efficiently plan hospitalization periods and plan personalized rehabilitation interventions based on the actual needs of the individual patient [31].

Unlike marker-based systems, markerless systems, especially those using a single camera, offer advantages such as lower costs, ease of use without the need for dedicated spaces and specialized personnel, and a less invasive experience for the patient, who does not need to wear markers. Moreover, these systems allow assessments to be performed in more natural contexts and are also suitable for uncooperative patients, such as newborns and young children [23].

Despite the many advantages offered by markerless motion capture systems, there are limitations related to their precision and accuracy, which are currently not yet fully comparable to marker-based systems, considered the gold standard in the field. This lower accuracy can be attributed to various technical and methodological factors, including [32, 30]:

- **Occlusion issues**: these occur when parts of the patient's body physically overlap with others, hindering the visibility of landmarks by Human Pose Estimation (HPE) algorithms and causing errors in their identification.
- **Technological limitations of the hardware used**: these include factors such as a narrow field of view of the camera, limited image resolution or a low frame rate. These characteristics directly affect the quality of the acquired data and, consequently, the accuracy of the reconstructed kinematic models.
- **Incorrect camera positioning**: A suboptimal camera alignment with respect to the anatomical planes of interest, such as the frontal plane and the sagittal plane, can significantly compromise the correct estimation of the orientation of the joint angles, especially in 2D analyses.
- **Influence of clothing**: The clothes worn by the subject during the motion acquisition could change shape dynamically, generating potential errors in the estimation of the position of the joint centers by the HPE algorithms
- **Use of suboptimal HPE models**: Each Human Pose Estimation model is typically trained using specific datasets, characterized by particular body

morphologies and types of movement. As a result, some movements and anatomical conformations may be detected with greater difficulty or less precision. This issue is further aggravated by the fact that most of the available open source datasets have not been specifically designed for clinical or biomechanical applications.

Despite the critical issues highlighted above, markerless systems represent the most suitable solution for telerehabilitation and exergame development, as they offer the best compromise between accuracy and practical feasibility. In fact, to perform an exergame, a very high precision in movement detection is not strictly necessary, but rather the use of cost-effective, accessible systems capable of operating in real time. However, given the limitation mentioned above, it remains essential to check these systems in systematic and rigorous validation, comparing them with the gold standards of reference. This process is essential for assessing the reliability of the kinematic parameters obtained via markerless systems and for evaluating their actual applicability in clinical and rehabilitative contexts.

2.2.1 Examples of markerless system used in exergaming and biomechanics analysis

Microsoft Kinect

One of the most widely used markerless systems for the development of exergames is certainly Microsoft Kinect, available in versions V1, V2 and Azure. Originally developed by Microsoft for gaming applications, the Kinect was subsequently embraced by the scientific community in various domains, including rehabilitation and biomechanical movement analysis. This success was mainly due to its integrated tracking algorithm which enables efficient real-time tracking, making it particularly well-suited for exergame development as well as for kinematic motion analysis and motor rehabilitation applications [33]. Across all its versions, Kinect integrates two types of cameras: an RGB camera and a depth camera, the latter used to measure depth, i.e., the distance between the camera and the objects or subjects in the field of view. Among its key strengths are its relatively low cost, portability, ease of use, and built-in tracking algorithm, which effectively exploits the information derived from depth. From a technical point of view, the first two versions of the Kinect use machine learning techniques based on Decision Forest to reconstruct a three-dimensional skeletal model of the analyzed subject [11, 34]. With the technological advancement of hardware and the introduction of increasingly powerful GPUs and CPUs, the latest version, the Azure Kinect DK, has integrated deep learning techniques, particularly convolutional neural networks (CNN), to enable highly accurate 3D tracking. This version of the Kinect is capable of identifying at least 28 body landmarks (up to a maximum of 33) for each subject, operating

in real time [35]. Thanks to these features and the robustness of its integrated algorithms, the Kinect is widely used in the development of exergames dedicated to the rehabilitation of patients affected by neurodegenerative diseases [21, 17]. Furthermore, the Azure Kinect version has also shown good performance in gait analysis applications, thus confirming its validity in the rehabilitation field [33].

OpenPose

OpenPose is an open source framework developed by Cao et al. [36, 37] designed to estimate the two-dimensional (2D) pose of multiple people within an image or video stream.

The algorithm is based on a multi-stage convolutional neural network (CNN) with two distinct branches: the first one generates confidence maps, which represent the probability that a given landmark (keypoint) is located in a particular pixel; the second produces the Part Affinity Fields (PAF), which are two-dimensional vectors that encode the position and orientation of the connections between the various parts of the human body. By combining these two informations, OpenPose allows a bottom-up approach to estimate the pose of a subject. This process initially involves the identification of individual body keypoints and subsequently their association with the corresponding person. This strategy allows the OpenPose algorithm to perform real-time tracking, identifying up to 25 landmarks for each subject, a feature that makes it well-suited for advanced human motion analysis. When deployed in hardware configurations involving multiple cameras, OpenPose is also able to obtain the three-dimensional (3D) position of the landmarks through triangulation techniques; in this case, however, tracking is limited to a single subject at a time. Using a hardware setup similar to the one adopted by the authors of the original framework (a laptop equipped with an NVIDIA GeForce GTX-1080 GPU), it is possible to achieve real-time performance. For instance, in the first version of the framework, the authors achieved a processing speed of approximately 8.8 frames per second (fps) when analyzing a video containing 19 individuals simultaneously. This result has been improved in the latest version of OpenPose, reaching a runtime of approximately 22 fps [37, 36, 25]. However, OpenPose has notable limitations in providing real-time performance when run on less powerful hardware, such as a common home laptop, which makes it unsuitable for the development of exergames [30]. Exergames, in fact, require immediate and continuous feedback to ensure smooth and responsive interaction with the user. They are also generally designed to be low-cost and accessible solutions, particularly suited to home rehabilitation settings.

AlphaPose

AlphaPose is a deep learning-based model developed for pose estimation, capable of estimating up to 17 two-dimensional (2D) landmarks per person. The model is designed to simultaneously identify multiple subjects within the same image, and has demonstrated excellent performance in terms of both accuracy and speed, as evidenced by the results obtained in standard benchmark datasets commonly used for evaluating Human Pose Estimation (HPE) models. AlphaPose can achieve real-time performance, provided it is run on sufficiently powerful hardware. For instance, the authors achieved their best performance using an NVIDIA 2080Ti GPU [38, 39]. AlphaPose currently does not provide a three-dimensional (3D) estimation of landmarks directly. However, it is possible to obtain this information using multiple properly calibrated cameras and specific triangulation algorithms to integrate information from different two-dimensional viewpoints, albeit at the cost of a more complex hardware setup [40].

To date, no exergames developed using this tracking model have been reported in the scientific literature. This absence could be due to the fact that AlphaPose does not guarantee real-time performance when run on standard hardware, not powerful enough, and does not natively offer three-dimensional information, thus limiting its use in applications that require immediacy and three-dimensional accuracy.

MoveNet

MoveNet is a model developed by Google using the Tensorflow platform, designed to perform fast and accurate two-dimensional (2D) tracking of 17 body landmarks. The model is optimized for real-time performance, with average speeds greater than 30 frames per second (fps), and is compatible with most commercial desktop, laptop, and mobile devices. MoveNet is available in two versions: Lightning and Thunderbolt. The Lightning version is lighter and less accurate, designed specifically for applications where low latency is a primary requirement. In contrast, the Thunderbolt version is slower but provides higher accuracy in estimating landmarks. The architecture of MoveNet is based on a MobileNet V2 network as a backbone, combined with four Feature Pyramid Network (FPN) layers [41, 42, 43]. MoveNet has already been applied in telerehabilitation contexts, integrated into systems aimed at remote monitoring of patients affected by musculoskeletal disorders. In particular, this model has been used to assess parameters such as strength, balance and range of motion during rehabilitation exercises [44].

Although MoveNet, especially in the Lightning version, has also demonstrated solid performance on mobile devices [45], there are currently no specific studies reporting its direct use in exergame development. This gap could be attributed to a lower accuracy compared to other HPE models and the lack of a direct 3D estimation. Therefore, obtaining depth information and 3D tracking would require

additional system configurations or algorithmic extensions.

MediaPipe Pose

MediaPipe Pose is a Machine Learning solution for pose estimation, part of the broader MediaPipe framework developed by Google. This framework allows the prototyping and implementation of models for image and video analysis, ensuring cross-platform compatibility such as IOS, Android and Python environments [46, 47]. This solution is capable of detecting up to 33 three-dimensional (3D) landmarks and generating a segmentation mask of the human body from a simple RGB input, using the **BlazePose** model, also developed by Google. MediaPipe Pose is notable for its high inference speed and can operate in real time even on mobile devices, achieving an average processing rate of 30 frames per second [48].

One of the key features of BlazePose, and by extension of Google MediaPipe Pose (GMP), is the ability to detect a higher number of landmarks (33) compared to the standard COCO topology, which provides only 17. The ability to track a greater number of landmarks, without significantly compromising processing speed, makes this model particularly suitable for detecting the position of specific body parts, such as hands, feet and face. To ensure computation speed, the model operates under the assumption that the face is always visible, so as to quickly detect it and obtain a region of interest (ROI) immediately usable to identify the complete tracking of the landmarks [49, 44]. GMP derives the third spatial coordinate (depth, z) using a variant of the BlazePose model that integrates information from the GHUM model. Initially, BlazePose outputs 33 two-dimensional coordinates; Google MediaPipe Pose then incorporates data from GHUM [50] to generate a personalized 3D model of the human body, adapted to the 2D landmark positions, thereby producing full 3D coordinates. This model is available in three different versions: lite, full and heavy, which are characterized by a progressive increase in accuracy at the expense of inference speed [51].

The model output is represented by a 33x4 array [52] which includes:

- **World coordinates (x, y, z)** : expressed in meters and referenced to an origin located at the center of the pelvis.
- **Visibility**: a value in the range [0.0, 1.0] that indicates the degree of visibility (confidence) of each landmark in the analyzed frame.

An example of the extracted tracking model is shown in Fig. 2.1.

Google MediaPipe Pose has been used as a pose estimation method in several biomedical analyses, such as posture [53] and gait analysis [54], due to its ease of use and low computational requirement. Its real-time processing capability, simplicity, and cross-platform compatibility make GMP particularly suitable for the development of exergames. Moreover, the ability to easily integrate these models

via Python, allows the data produced by the system to be used in popular game engines, such as Unity, thus facilitating the creation of interactive and functional exergames. An additional strength lies in the specific predisposition of GMP to single-person tracking, although multi-person tracking is also supported. For these reasons, Google MediaPipe Pose was selected as the pose estimation model for the exergame used in this thesis.

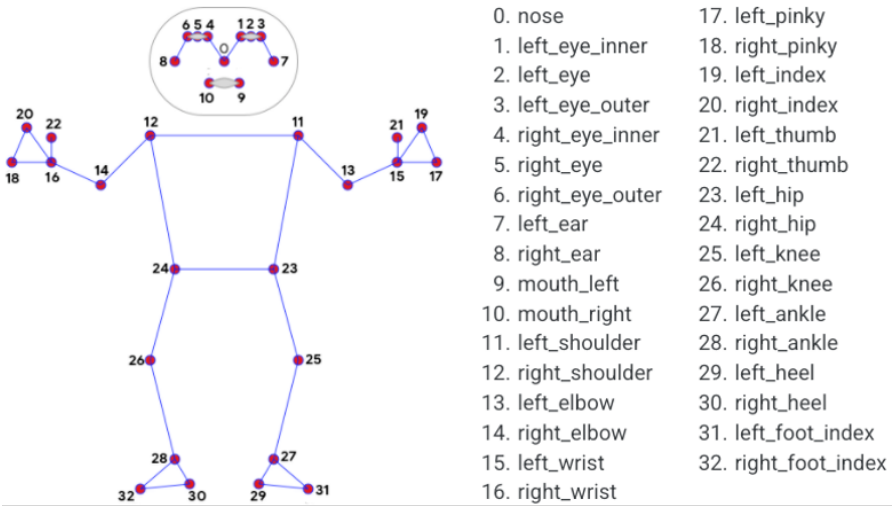

Figure 2.1: MediaPipe Pose landmarks [52]

2.2.2 The importance of validation

To obtain clinically relevant information, especially in the context of rehabilitation, it is essential to have reliable and sufficiently accurate data. Telerehabilitation has the significant advantage of allowing, through specific tools such as exergames, continuous monitoring of patients and constant acquisition of objective data, thus reducing the need for frequent in-person assessments. In the specific case of the exergame used in this study, the goal is not to achieve the extremely high precision typical of specialized motion analysis laboratories, but rather to provide meaningful support to the rehabilitation process by capturing basic kinematic data. For example, if a given rehabilitation exercise requires a minimum arm flexion of 60°, the system must be able to produce a value close to this reference in order for the assessment to be considered useful, reliable and clinically relevant. Validation against an accurate reference system (gold standard) allows to quantify the extent of the measure differences and to determine the level of reliability of the data produced by the exergame to perform functional assessments.

Several studies have addressed this issue. Amprimo et al. [28] validated the

Google MediaPipe Hand (GMH) hand tracking system, both in the standard version and in an enhanced version developed by the same authors (GHM-D). The study highlighted strong temporal and spectral coherence with respect to gold standard systems, as well as greater accuracy in the optimized version. These findings demonstrated that the enhanced version is also applicable to clinical assessments, in particular in the functional assessment of the fine hand movements in subjects with neurological disorders.

Dill et al. [55] investigated the accuracy of the MediaPipe Pose system during the execution of common physical exercises, such as squats and push-ups. The results show that the accuracy of GMP is strongly influenced by the camera angle with respect to the subject; the system provides reliable data in optimal record conditions, while deviations from these conditions significantly increase the error.

Lafayette et al. [56] assessed the accuracy of joint angle estimation using both RGB-D and RGB cameras in combination with GMP. The authors examined simple exercises, such as knee flexion and extension, as well as shoulder flexion, extension, abduction, and adduction. Results indicated that the exclusive use of RGB cameras yielded joint angle estimates more consistent with gold standard measurements, whereas RGB-D configurations produced less satisfactory results.

In conclusion, validation represents an essential step to fully understand the potential and limits of the adopted solution. This process is particularly important in the rehabilitation field, where consistent and reliable data are essential to effectively support the clinical decision-making process.

Chapter 3

Methods

3.1 "Palestra": An Exergame for Upper Limb Rehabilitation

The following paragraph aims to provide a general description of the exergame employed in this thesis project. The previous prototype of the "Palestra" exergame was enhanced by the CNR-IEIIT and was designed to support upper limb mobilization and enable quantitative movement analysis, in order to detect anomalies, improvements, or deteriorations in the patient's motor performance resulting from the functional decline associated with Parkinson's disease and ageing.

Hardware and Software Requirements

The video game was developed using the Unity framework (editor version 2021.3.18) and programmed in C# (version 7.0). Regarding the hardware requirements, the game was initially developed and tested on a desktop computer with the following technical characteristics:

- Operating system: Windows 10;
- Processor: Intel® Core™ i7-10700 CPU, 2.90 GHz;
- RAM: 32 GB;
- GPU: NVIDIA GeForce RTX 2070 Super;
- External camera with a resolution of 1920x1080 pixels, 30 fps;

To verify the correct functioning and performance on other devices, the game was subsequently installed on a laptop computer with similar characteristics, except for the operating system (Windows 11) and the camera used (integrated camera

with a resolution of 3840x2160 pixels, 60 fps). This latter setup was used to acquire the data employed in the subsequent analyses, as described in Section 3.4.

The tests carried out on both hardware configurations did not reveal any noticeable latency during gameplay, thus satisfying the requirement of real-time motion capture. This observation is consistent with findings in the literature, which report that exergames are generally tolerant to latencies in the range of several hundred milliseconds without compromising the user experience. [57] Regarding software, the exergame is composed of two main modules: one developed within the Unity environment, where the scenes and the logic for game management and control were implemented, and another developed in Python, dedicated to motion tracking using the Google MediaPipe Pose library (version 0.10.13, see Section 2.2.1). The two modules communicate through the pipeline illustrated in Figure 3.1

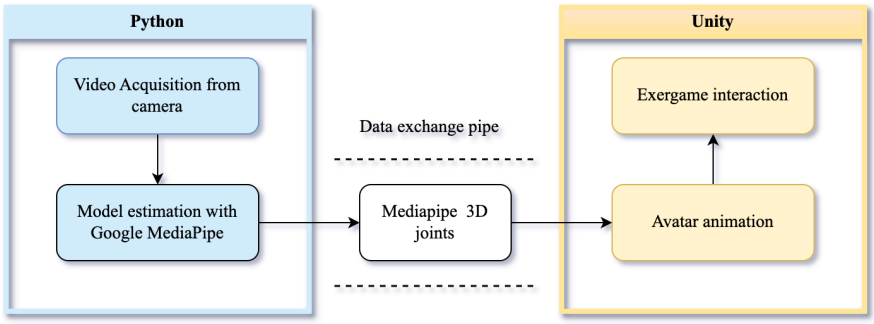

Figure 3.1: Communication pipeline between Python and Unity.

As shown in Figure 3.1, the Python component acts as the **server**, while the Unity component acts as the **client**. Specifically, the Python module acquires and processes the video stream from the camera, extracting from each RGB frame the two-dimensional positions of 33 anatomical reference points (*landmarks*), as illustrated in Figure 3.2. These positions, expressed as 2D coordinates normalized to the image resolution, are then used to derive the three-dimensional positions (*world landmarks*) of each point, relative to the center of gravity of the body, approximately located at the midpoint between the hip landmarks in the pelvic region. These three-dimensional coordinates are absolute, independent of the image resolution, and expressed in *meters*. Although the x and y coordinates have good accuracy, as they are located on the image plane, **the z coordinate, representing the depth, is only an estimation made by the model.**

The adopted reference system is defined as follows:

- The x coordinate increases toward the right;

- The y coordinate increases upward;
- The z coordinate is positive when the landmark is located in front of the plane containing the body's center of gravity, and negative when it is behind.

All the 33 landmarks are estimated for each frame, then packaged and transmitted to the Unity module, which uses them to animate the game avatar in real time based on the user's movements. To optimize animation performance, some less accurate landmarks, such as those corresponding to the hands and feet, are excluded. Consequently, the avatar animation of the upper and lower limbs is limited to the wrist and ankle landmarks, respectively. For facial animation, a single landmark is used, calculated as the midpoint of all detected landmarks of the face.

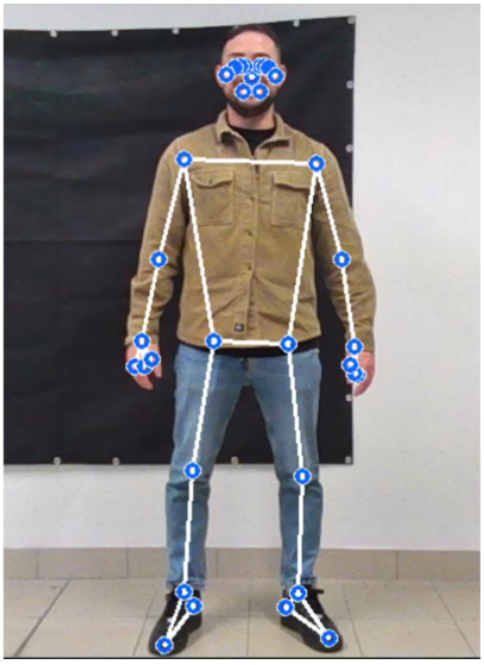

Figure 3.2: Landmarks extracted by Google MediaPipe.

Description of the Game Environment

The Unity module manages the game environment and enables interaction with the elements contained within it. The exergame was developed to be controlled through simple upper limb movements performed by the subject under evaluation. However, for user registration and for the configuration of accessibility settings, such as the impaired limb, game mode, and difficulty level, the assistance of an external operator, typically a caregiver is required. The game interface was designed to be simple and intuitive, specifically tailored for users with little or no prior experience with video games or technology in general. Since the main goal of the game is to promote upper limb mobilization, the interface is deliberately kept free of unnecessary elements, thus minimizing potential distractions.

From an implementation perspective, the exergame is mainly structured around a central scene named *Palestra*, in which the animated avatar is displayed based on the three-dimensional landmarks received from the Python module. This main scene was designed to be consistent with the exercises to be performed, recreating a gym-like environment using free assets available on the Unity store, such as dumbbells, mats, and generic fitness equipment (see Figure 3.3). The central area, intended for avatar placement, is deliberately kept free of additional elements to avoid visual interference with the execution of the exercise. In addition to the main scene, the game includes secondary scenes in the form of semi-transparent panels, which allow the user to log in and configure various gameplay modes, while keeping the main scene visible in the background.

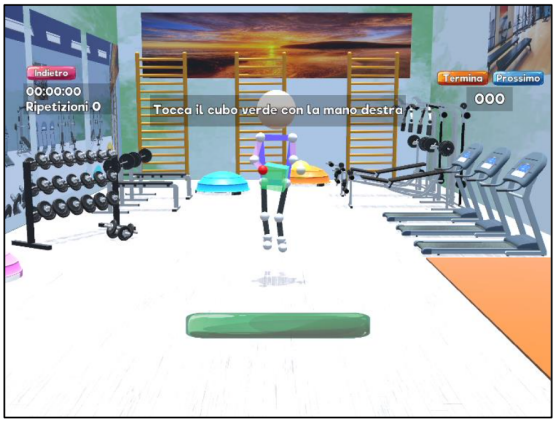

Figure 3.3: Main scene "*Palestra*" with the Avatar inside.

As previously described, the main scene features the avatar, which replicates the user's movements in real time. To prevent delays in graphical rendering, an avatar made of simple three-dimensional elements (spheres and lines) representing joints and limbs was chosen, simulating a simplified version of the human skeleton (see

Figure 3.3). In total, the avatar consists of 15 spheres, each linked to a corresponding three-dimensional landmark received from the Python module, enabling real-time animation. The head is represented by the largest sphere and is animated using a landmark obtained as the midpoint between those detected on the user's face. To avoid confusion between the right and left sides, the avatar is mirrored relative to the player. Furthermore, the spheres representing the hands, connected to the wrist landmarks, are color-coded: blue for the left hand and red for the right hand.

Figure 3.4 illustrates the configuration panel. This setup, usually performed by the operator or caregiver, allows the user to specify the following options:

- **Interaction hand**: This item indicates the hand the patient will use to interact within the game scenes, for example to initiate an exercise. This typically corresponds to the less impaired limb.
- **Difficulty level**: This item defines the minimum angle that the limb must reach to register a complete movement. Since different levels of impairment may be present between the arms, difficulty levels can be configured separately.
- **Game mode**: the exergame offers two operational modes:
  - *Timed*: This mode defines a time window (in seconds) during which the subject must perform as many repetitions as possible. This setting does not differentiate between arms or exercise type.
  - *Repetitions (default)*: This mode defines the minimum number of repetitions for each exercise to be completed within a time window defined as 3 seconds per repetition. For example if 5 repetitions are selected, the exercise will last 15 seconds.
- **Control mode**: two control options are available:
  - *Automatic control (default)*: This mode allows the subject to move his arms freely without specific constraints.
  - *Manual control*: In this mode an operator determines when and which movement should be performed, using visual cues. This mode is particularly useful for assessing the cognitive function of the subject.

The main scene features a graphical user interface (GUI) designed to assist the patient through several elements:

- **Central textual elements**, including an upper section (for immediate instructions) and a lower section (to direct attention during the exercise).
- **Lateral text elements**, which display auxiliary information (e.g., on the left: exercise duration and repetition count; on the right: partial score).

- **Temporary textual elements**, such as the initial countdown and the final score.
- **Auditory cues**, essential for providing immediate feedback on the execution of the exercises.
- **Visual cues**, such as colored arrows indicating which arm to move, are particularly useful in simultaneous or alternating modes or when the game is set to manual mode.
- **Interactive elements**, such as the green cube used to initiate the countdown and start the exercise. If the user is unable to activate it autonomously, the operator can start the session manually.
- **Operator controls**, including buttons to interrupt the session, skip an exercise, or return to the configuration screen.

Figure 3.4: Configuration panel.

3.1.1 Types of Proposed Movements, and Motor and Cognitive Aspects

The exergame was primarily developed to stimulate motor functions through repetitive movements of the upper limbs. In addition to promoting joint mobility, it was also designed to train motor control and coordination. Specifically, the exergame includes the following exercises:

- **Frontal arm raises:** starting from an upright resting position with the arms extended alongside the body, the user raises the arms forward and then returns them to the starting position.

- **Lateral arm raises**: also starting from an upright resting position with the arms at the sides, the user raises the arms laterally and then returns them to the initial position.

Each exercise can be performed in one of four execution modes:

- **Single right arm**: the user raises only the right arm.
- **Single left arm**: the user raises only the left arm.
- **Simultaneous**: the user raises both arms at the same time.
- **Alternating**: the user raises the right and left arms alternately.

Single arm raises are primarily designed to stimulate control and mobility of the affected limb, while simultaneous and alternating raises are also designed to enhance motor coordination.

As anticipated, the "manual" control mode, managed by the operator, provides an opportunity to further stimulate the cognitive functions of the patient. In this mode, the patient must perform specific movements indicated by visual cues represented by arrows that appear in the main scene. The operator decides not only when to activate the visual cues, but also, in the alternating mode, which cue (right or left) to display. Therefore, this mode enables the evaluation, for example, of the patient's reaction times, i.e. the time span between the visual cue and the start of the corresponding movement by the patient. In this way, in addition to motor stimulation, attention, planning and problem solving skills are also stimulated, which are the distinctive elements of cognitive training.

A key feature of exergames is the reward system associated with the correct performance of the assigned tasks: for each correctly performed movement, the user earns 100 points, while for each incorrectly performed movement, 25 points are deducted. The user's final score corresponds to the sum of the points obtained for each individual movement, and at the end of the exercise a panel with the overall score is displayed (see Figure 3.5).

The reward system represents a fundamental component in the proper design and implementation of an exergame, as it can enhance various aspects associated with the patient's performance, in particular [58]:

- **Motivation**: the reward system enables the user to experience a sense of personal satisfaction and achievement, promoting a stronger inclination to continue the activity.
- **Reinforcement**: rewards encourage the user to explore and adopt new strategies and approaches to the assigned task, thereby increasing the probability of achieving the predefined therapeutic goals.

Figure 3.5: Scoring panel.

- **Feedback**: rewards serve as immediate feedback, helping the user understand whether the exercise is being performed correctly, thus promoting learning and self-correction.
- **Engagement**: the use of reward strategies increases the level of user engagement, encouraging active and consistent participation, and, consequently, amplifying the therapeutic benefits achieved through the exergame.

3.1.2 Output

The exergame generates two output files in JSON format: the first contains all information related to the skeletal model obtained through Google MediaPipe Pose, while the second contains summary data regarding the exercise performed. Data recording begins when the game countdown starts, either through interaction with the green cube by the patient or through direct activation by the operator. The initial 3 seconds of the countdown, during which the patient maintains a static position, are also included in the recording, as this data may be useful for subsequent analysis of postural features. Recording ends upon completion of the exercise or when the time limit expires.

Specifically, the file related to the skeletal model includes:

- **Normalized coordinates**: x, y and z coordinates derived from the position of the landmarks in the video frame. These coordinates, normalized with respect to the resolution of the camera used, are strictly dependent on the hardware employed.
- **World landmarks**: absolute coordinates, expressed in meters, of the landmarks relative to a reference system centered at the pelvis.

- **Confidence**: a coefficient between 0 and 1 indicating the reliability of the estimated position of each landmark.
- **Timestamp**: the time assigned to each acquired frame, measured from the beginning of the recording.
- **FrameID**: a progressive identifier assigned to each frame processed by the Python component, useful for detecting any frame loss during acquisition.

The file containing the summary information of the exercise includes:

- Patient ID and Session ID: unique identifiers for the test performed.
- Start and end time of the exercise.
- Exercise ID:
  - 0: single right arm.
  - 1: single left arm.
  - 2: alternating execution.
  - 3: simultaneous execution.
- Movement plane: frontal or lateral.
- Selected difficulty level.
- Game time: time taken to complete the exercise.
- Number of correct movements and errors.
- Exercise completion indicator.
- Operator interruption indicator.
- Time expiration indicator, in case the exercise is interrupted due to exceeding the predefined time limit.
- Involved body part: right, left, alternating, or both.

These JSON files, particularly the one related to the skeletal model, serve as the basis for the analyses conducted in this thesis project.

3.2 Automated pipeline for joint signal analysis

The aim of the analysis is to extract joint angle signals from the landmark trajectories obtained through the exergame. In order to generate high-quality signals, suitable for deriving parameters that can provide an accurate assessment of motor task execution, an automatic purpose-built processing pipeline has been developed.

This pipeline is specifically optimized for processing the data collected by the game and has been designed to allow the extraction of both angular signals and outcome parameters. The analysis was performed using MATLAB R2024b, chosen for its flexibility in numerical computing and its extensive library of built-in functions.

The input data used are the JSON files and include the estimated positions of the body landmarks over time, along with a corresponding timestamp vector (see Section 3.1.2). The developed pipeline includes the following main steps: data loading, application of a moving median filter, resampling, low-pass filtering, computation of the 3D vectors representing the selected body segments, joint angle calculation, signal segmentation and extraction of the outcome parameters. The entire pipeline process is schematically illustrated in Figure 3.6.

Each processing step is applied to all signals; however, specific differences exist regarding algorithm parameters used in the signal pre-processing phase, such as the size of the moving median filter window. Additional differences arise during the feature extraction and the outcome computation stages, depending on the type of signal being analyzed. In particular, bilateral arm raise tasks require different segmentation strategies and distinct outcome metrics to be computed. The choice of employing different parameter values is justified, for example, by the qualitative differences observed between frontal and lateral movement acquisitions. In particular:

- Lateral arm raises produce higher-quality signals, as movement occurs mainly in the plane parallel to the camera's point of view.

- Frontal arm raises, on the other hand, exhibit lower signal quality and tend to be more affected by impulsive noise, as the movement occurs along the plane perpendicular to the camera (z direction). Since this component is estimated, it is more susceptible to errors, artifacts and noise, further accentuated by self-occlusion phenomena.

Due to the significant differences in the signal quality between the two types of executions, it was necessary to adapt specific parameters in the parameter estimation algorithms, which will be described in detail in the following paragraphs.

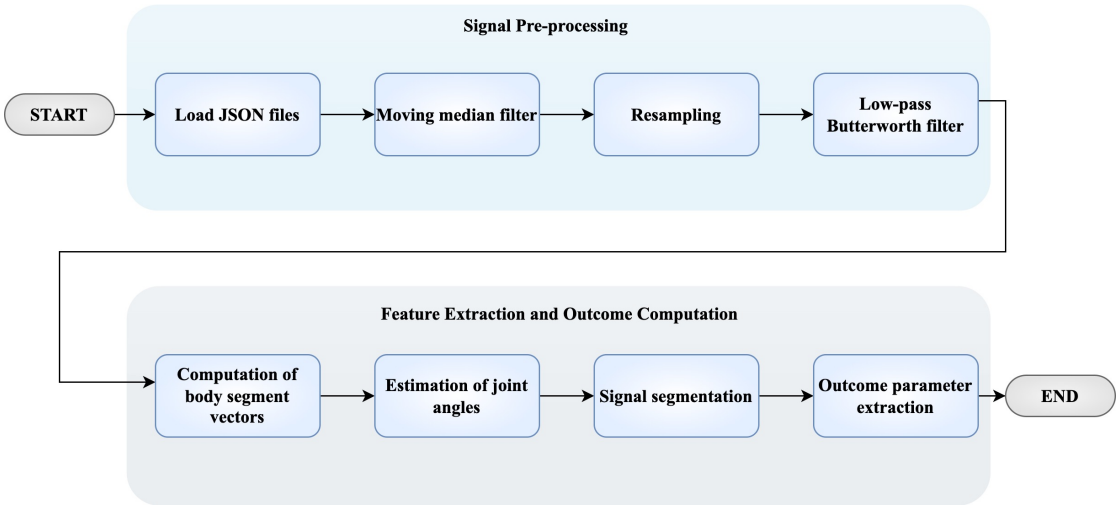

Figure 3.6: Signal processing pipeline.

Load JSON files

In the first step, the JSON files generated by the exergame are loaded using the JSONLab toolbox [59]. To quickly identify any potential issues in the acquired data, basic information is computed to the screen, including the overall duration of the exercise (in seconds), the total number of recorded frames, and the average, minimum, and maximum frame rate. Regarding joint angle calculation, not all the data contained in the JSON files are necessary. Therefore, a refined data structure is extracted, containing only the following data: *world landmarks*, *visibility* and *timestamp*. Visibility information was crucial in the design of the pre-processing algorithms and in identifying the technical limits of the exergame. However, it was not integrated as a final clinical outcome parameter, given its limited clinical relevance.

An example of angular trajectories corresponding to the single arm raise exercise, in lateral and frontal directions, is shown in Figure 3.7.

As shown in the graphs, there is a significant difference in spatial coordinates between the lateral and frontal tasks. In lateral raises, the movement occurs primarily in the plane parallel to the camera view; as a result, the x and y coordinates exhibit the largest variations. In contrast, the z coordinate has smaller fluctuations, which are likely attributable to minimal movements along z direction. In frontal raises, however, the z coordinate shows a greater excursion than the x coordinate, as the movement occurs along the axis perpendicular to the camera plane. The associated signal is more irregular than in the lateral movements, often containing spikes caused by occlusion and suboptimal wrist positioning relative to the camera. These conditions make it difficult for MediaPipe Pose to perform

an accurate estimation of z (depth) component; for this reason, it is essential to consider this factor to adequate data pre-processing.

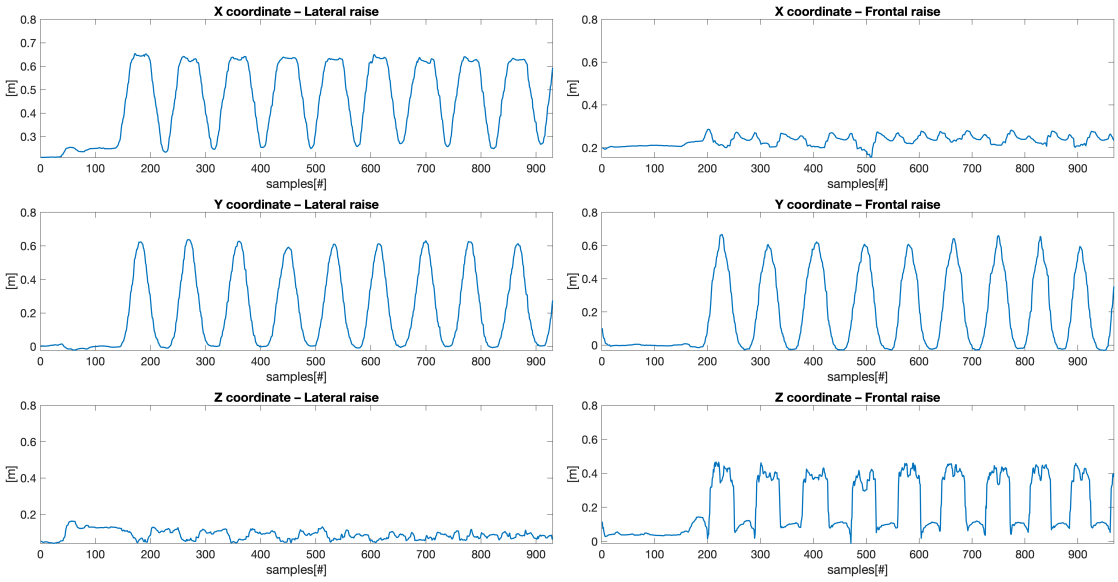

Figure 3.7: Comparison of raw wrist trajectories in lateral and frontal left arm raises.

Moving median filter

To reduce impulsive noise, typically characterized by isolated peaks (spikes), especially in the z-coordinate, a moving-window median filter was implemented using the MATLAB built-in function *medfilt1*. The use of the median filter is motivated by the hypothesis that random signal fluctuations can be interpreted as outliers, i.e. extreme values isolated from the surrounding data points. Since the median is known to be less sensitive to outliers than the mean, the median filter proved to be effective in removing spikes. An additional advantage of the median filter is that, in signals where there are no significant impulsive fluctuations, such as in the case of the x- and y-coordinates, it produces only a minimal smoothing effect, comparable to that of a moving average filter. However, the latter approach is not effective for the z-coordinate, due to its high sensitivity to outliers

For the previous mentioned differences in signal quality, it was necessary to tailor the size of the moving window of the filter to the specific type of signal processed. This choice was made to avoid excessive signal smoothing, which would compromise the performance of the segmentation algorithm and, consequently, hinder the accurate extraction of temporal parameters. These parameters are, in fact, strictly dependent on the morphological structure of the segmented signal. Therefore, for the lateral movement tasks, whether involving a single arm or both

arms, a window size of 8 samples was chosen. In contrast, for all tasks that involve frontal arm raises, a larger window size of 16 samples was used.

Resampling

The signal produced by the exergame is not characterized by a constant frame rate. The analysis of the timestamps revealed significant variability in the instantaneous frame rate, ranging from 25 to 40 frames per second (fps), with an average generally falling between 30 and 34 fps. This variability is mainly influenced by the hardware characteristics of the system on which the exergame is executed, as well as by the instantaneous computational load of the CPU and graphic cards. For example, if the system runs more applications simultaneously, the average frame rate tends to decrease. Various operating conditions were therefore tested, such as running background applications or using the laptop without an active power source, finding a decrease in the average frame rate, which, however always was greater than or equal to 30 fps.

To standardize the sampling frequency, the signals were resampled at 50 Hz using linear interpolation, implemented via MATLAB's built-in function *interp1*. The decision to resample the data at 50 Hz was motivated by the following factors:

- This frequency is sufficiently close to the observed average frame rate (~35 fps), thereby minimizing the extent of interpolation applied to the signals;
- Resampling at high frequency would have unnecessarily increased the size of the signals, causing potential artifacts and a general slowdown in the execution of subsequent processing algorithms;
- A very high temporal resolution is not required to extract clinically relevant kinematic and temporal parameters from the acquired signals.

The use of linear interpolation proved effective, as the overall morphology of the signals remained largely unaltered in all the cases analyzed.

Low-pass Butterworth filter

To further attenuate the residual high-frequency noise, the signal bandwidth was limited to 10 Hz, applying a fourth-order Butterworth low-pass filter. The optimal filter order was determined using the built-in MATLAB function *buttord*, while the filtering itself was performed using the *butter* function, also available within the MATLAB environment.

Following the application of the moving-window median filter, resampling to 50 Hz and subsequent low-pass filtering, the resulting trajectories exhibit a significant reduction in noise, while preserving their essential morphological features.

To visualize the effect of signal pre-processing phase, the raw trajectory was resampled to match the time baseline of the filtered signal. This allows a direct comparison of the signal quality before and after the complete pre-processing block (median filtering, resampling, and low-pass filtering), as shown in Figure 3.8.

As shown in the figure, the z-coordinate exhibits a smoother morphology after the preprocessing steps, expecially in frontal movements while the other coordinates remain approximately unchanged.

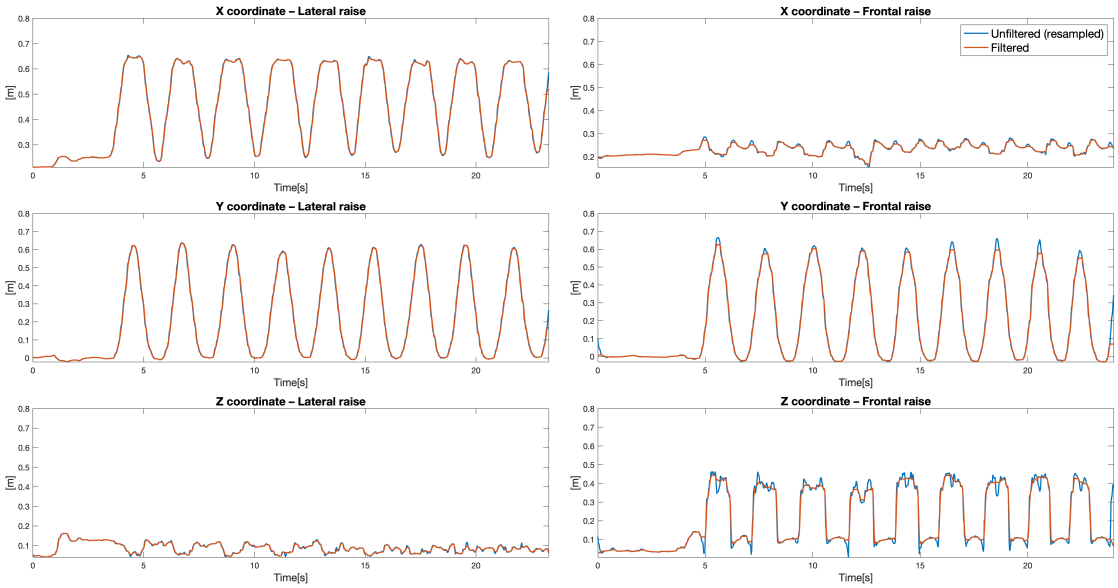

Figure 3.8: Comparison between the resampled unfiltered signal (blue line) and the signal (red line) after preprocessing (median filtering, resampling, and low-pass filtering). Resampling was applied to the raw signal solely to align the time axes, allowing for direct visual comparison.

Computation of body segment vectors

In order to define and compute a joint angle, it is necessary to first define the two vectors between which the angle is measured. Each vector is calculated as the difference between the spatial coordinates of the two anatomical landmarks defining the body segment of interest. Table 3.1 presents the vectors considered in the analysis, along with the corresponding landmarks used for their computation.

Vector name	Landmarks used for vector computation
Right arm	Right shoulder (11) – Right wrist (15)
Right trunk	Right shoulder (11) – Right hip (23)
Left arm	Left shoulder (12) – Left wrist (16)
Left trunk	Left shoulder (12) – Left hip (24)

Table 3.1: Identified vectors and corresponding anatomical landmarks used for angle computation. Numbers in brackets refer to the landmark indices as shown in Figure 2.1.

Estimation of joint angles

Once the vectors have been defined, joint angles are computed, expressed in degrees, using the following formula:

$$\phi = \arctan\left(\frac{|v_1 \times v_2|}{v_1 \cdot v_2}\right)$$

Where $|v_1 \times v_2|$ indicates the magnitude of the cross product between the two vectors defining the joint angle (e.g., "right arm" and "right trunk"), and $v_1 \cdot v_2$ represents their scalar (dot) product.

Figures 3.9, 3.10 and 3.11 show some examples of angular trajectories for both lateral and frontal raises, performed with the right and left upper limbs.

As can be observed, all the trajectories relating to the lateral raises exhibit less irregularities than the frontal ones (Figure 3.9), despite the latter were processed using a median filter with a wider moving window. The most pronunced oscillations tend to occur near the minimum and maximum points of the trajectories. This is likely due to the relative stillness of the landmark during those intervals, making the estimation of the z-component particularly sensitive and prone to fluctuation; clearly observable also in the positional (meter-based) trajectories from which the joint angles are derived.

In alternating raises (Figure 3.10), a transitory oscillation between two consecutive repetitions can be observed, which is absent in the simultaneous task. This deformation seems to be due to pose estimation inaccuracies that arise when one limb is in motion while the other remains temporally stationary. This undulation tends to be convex in the lateral task and concave in the frontal one.

The choice of the window size for the median filter was also made taking into account this artifact; in fact, an excessive smoothing would shift the minimum of a repetition too far from its corresponding peak, resulting in overestimated temporal parameters during subsequent segmentation.

During simultaneous movements (Figure 3.11), the tracking is generally more stable, likely because coordinated bilateral movement facilitates more accurate estimation of the limbs' 3D positions.

Although the exergame was configured to execute 10 repetitions per exercise, only 9 were saved, as the record ended when the angular threshold of the last repetition was reached. In the specific case of alternating raises, the 10 planned repetitions are equally divided between the two limbs (5 repetitions per limb), but again, one repetition was not captured, leading to only 4 recorded repetitions for the left limb. However, this behavior did not significantly impact the subsequent analysis, as the absence of a single repetition did not compromise the extraction of representative parameters.

In frontal raises, the system sporadically misidentifies a single movement as multiple repetitions, thus reducing the actual number of repetitions registered. For example, even when 10 repetitions were configured, sometimes a reduced number of repetitions may appear in the output: one lost due to the saving issue and the other miscounted, causing the task to end prematurely. Nevertheless, these sporadic irregularities did not affect the integrity of the analysis, as the extracted parameters remained consistent and representative across all subjects.

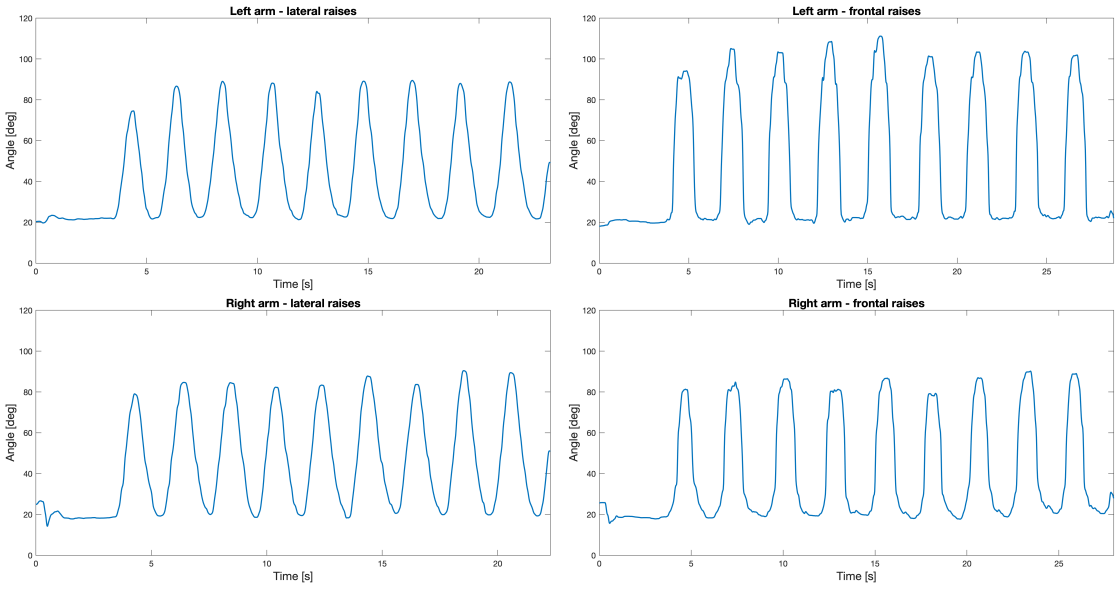

Figure 3.9: Angular trajectories during single-arm lateral and frontal raises. The trajectories obtained during lateral raises are characterized by a more regular morphological profile compared to those from the frontal raises, likely due to the more favorable tracking conditions when the movement occurs within the plane of the camera.

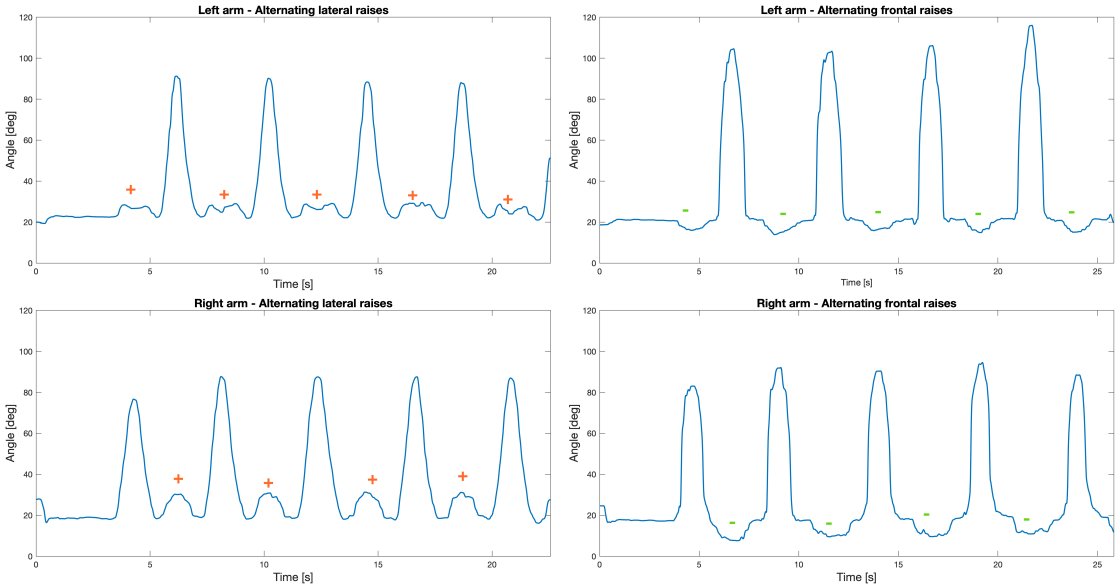

Figure 3.10: Angular trajectories during alternating raises. Alternating lateral raises are characterized by convex transient oscillations (+) between repetitions, while alternating frontal raises exhibit concave (-) ones. These artifacts are likely related to pose estimation inaccuracies during asymmetric arm movement.

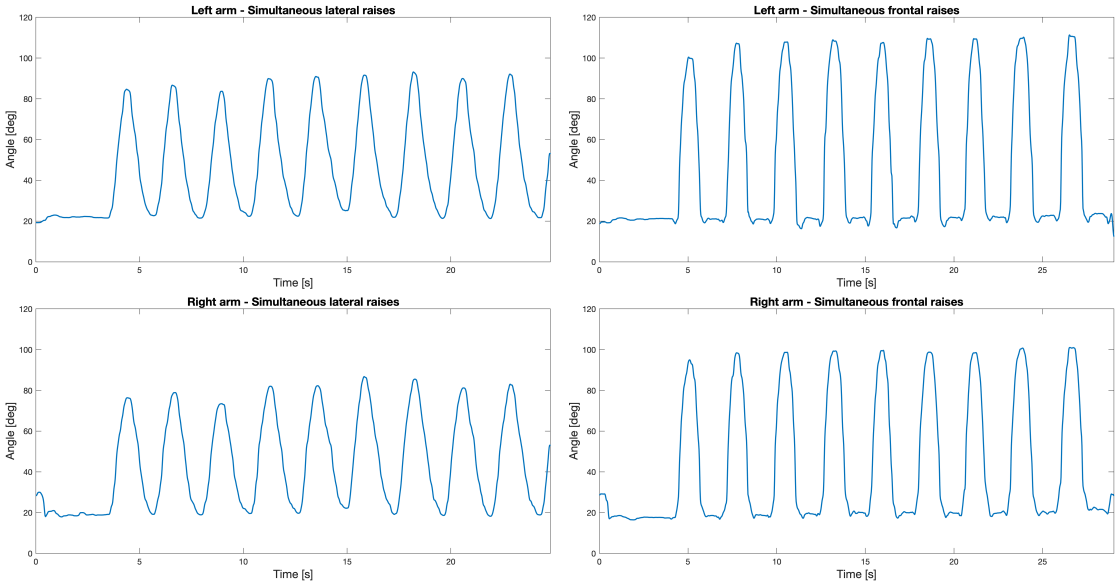

Figure 3.11: Angular trajectories during simultaneous raises. Simultaneous raises are characterized by smoother profiles compared to alternating raises, likely due to more stable pose tracking during symmetric arm movements.

Signal segmentation

Starting from angular trajectories, individual repetitions are isolated to extract representative and descriptive outcome parameters of the performed movement. Signal segmentation is carried out by identifying characteristic points in the trajectory, specifically the local maxima and minima.

Maxima detection

In the first stage, signal peaks (local maxima) are identified using the MATLAB built-in function *findpeaks*. To ensure an accurate selection, i.e., one maximum per repetition, it is necessary to properly set the function's parameters: *MinPeakDistance*, *MinPeakHeight* and *MinPeakProminence*.

To obtain an initial estimate of these parameters, *findpeaks* function is initially applied, by setting *MinPeakHeight* to the mean of the signal plus 80% of its standard deviation. In this way, three distinct vectors are obtained: the first containing the value of the identified maxima, the second relating to their temporal position (expressed in samples), and the third containing the prominence (a measure of how much each peak stands out from its surroundings, computed as the vertical distance between the peak and its lower neighboring minima) [60].

In the second stage, *findpeaks* function is applied again, this time with all final parameters properly set, as summarized in Table 3.2. The complete procedure for final peak identification consists of the following steps:

1. Compute *MinPeakHeight* as:

$$MinPeakHeight = \text{mean}(signal) + 0.8 \cdot \text{std}(signal)$$

2. Initial application of *findpeaks* function using only *MinPeakHeight*, obtaining a first provisional set of maxima (value, position, and prominence).

3. Estimate *MinPeakDistance* as the average number of samples between consecutive provisional maxima, computed from the differences in their positional indices.

4. Final application of *findpeaks* using all three parameters:*MinPeakHeight*, *MinPeakDistance*, and *MinPeakProminence* (as detailed in Table 3.2), producing a refined set of maxima while minimizing false positives.

Parameter	Computation	Definition
MinPeakHeight	$\text{mean}(signal) + 0.8 \cdot \text{std}(signal)$	Minimum height that a peak must exceed to be considered valid
MinPeakDistance	$0.8 \cdot \text{PeakDistance}$	Minimum distance (in samples) required between consecutive peaks
MinPeakProminence	$0.5 \cdot \max(\text{prominence})$	Minimum prominence value required for a peak to be considered valid

Table 3.2: Parameters used in *findpeaks* function.

Minima detection

Once the signal maxima have been identified, a custom algorithm, referred to as *Ultimate_min*, has been implemented to detect the corresponding local minima. This algorithm allows to identify two minima for each previously identified maximum: one preceding (on the left) and one following (on the right).

This procedure requires the following inputs:

- the angular trajectory of interest;
- the vectors containing the values and positions of the previously identified maxima;
- a spatial threshold value chosen properly (default value: 40°).

The algorithm returns two vectors as output: one containing the indices (positions) of the detected minima, and the other containing their corresponding angular values.

The objective of the algorithm is to select two minima for each maximum, respecting two fundamental constraints:

- **Spatial constraint**: the angular value of the selected minimum must be less than or equal to the defined threshold (default: 40°);
- **Temporal constraint**: the selected minimum must be temporally close to the reference maximum. Specifically, the distance must not exceed 40 samples (equivalent to approximately one second, considering a sampling frequency of 50 Hz).

The search for minima is performed using the built-in *findpeaks* function. Since this function is designed to identify local maxima, the angular signal is first inverted to enable the correct identification of the minima.

The detailed procedure includes the following steps:

1. **Signal inversion**: the angular signal is inverted. A preliminary threshold is then computed as the mean of the inverted signal and used as the *MinPeakHeight* parameter in the *findpeaks* function. This step yields a provisional set of local minima.

2. **Preliminary candidate selection**: for each previously identified maximum, the two closest minima are selected, one to the left and one to the right.

3. **Candidate evaluation**:

   - The first candidate (the closest minimum) is evaluated exclusively on the basis of the spatial constraint (angular threshold). If it satisfies this condition, it is immediately selected.
   - If the first candidate fails the spatial constraint, the second candidate is considered. This one is evaluated against both the spatial and temporal constraints.
   - If the second candidate satisfies the spatial constraint but not the temporal one, it is discarded and the first candidate is selected as a fallback solution.

4. **Iteration**: the procedure described is repeated for each maximum present in the signal.

This approach allows accurate and robust detection of local minima, minimizing the likelihood of selecting spurious points due to signal noise or random fluctuations. It is important to note that a single minimum may be shared between two adjacent maxima (e.g., the right minimum of one peak may coincide with the left minimum of the next).

The application of this segmentation procedure is shown in Figure 3.12.

At the end of this process, four vectors are obtained, containing the positions and corresponding values of all maxima and minima, which serve as the basis for the subsequent computation of the outcome parameters.

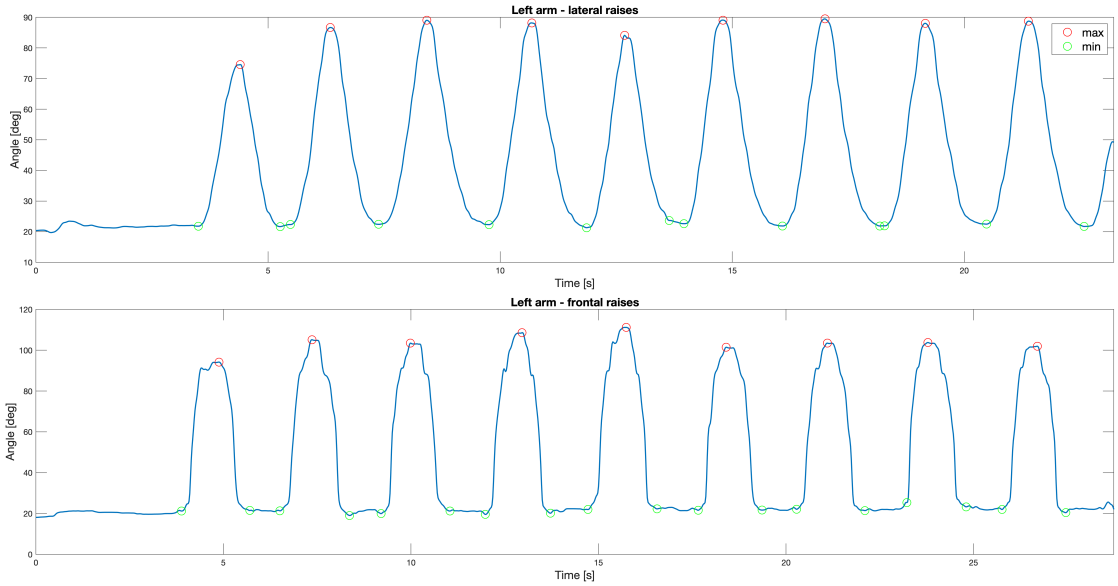

Figure 3.12: Angular signal segmentation based on the detection of local maxima and minima. The figure shows the shoulder joint angle trajectories during lateral and frontal single-arm raises, with the peaks (red) and minima (green) used to delimit single repetitions.

3.3 Outcome parameters

In order to evaluate the motor performance of subjects who engaged with the exergame, several kinematic parameters were computed from the previously obtained angular trajectories. Each parameter was selected to capture a specific characteristic of the movement, covering different aspects such as temporal information, spectral content, spatial descriptors, and metrics related to velocity and motor coordination.

Most of the kinematic parameters were extracted from the segmented signals. However, some specific metrics, such as the cross-correlation lag in bilateral movements, were computed directly from the complete angular trajectory. These parameters are not averaged across multiple repetitions, as they are not based on segmentation. Nevertheless, these metrics provide useful functional information and are therefore included in the overall assessment of the movement.

Generally, segmentation-based parameters were averaged across individual repetitions in order to minimize variability introduced by the segmentation process and to reduce the impact of transient signal fluctuations. These average values are especially relevant for the clinical evaluation of motor function.

Moreover, certain parameters were calculated for each individual repetition. Although these are not directly used for clinical evaluation, they were retained for validation purposes, enabling the construction of a broader and more consistent

dataset.

The outcome parameters described are common across all motor tasks, with the exception of specific indices used exclusively in bilateral exercises to assess inter-limb coordination.

In the following paragraphs, each outcome parameter will be described in detail, grouped into two categories: the first related to unilateral-arm tasks, and the second concerning bilateral exercises. A summary list of all parameters is provided in Tables 3.3 and 3.4 to support quick reference and comparison.

Domain	Parameter	Unit	Description
Angular	High Amplitude (HA)	deg	Mean and SD of maximum angles
	Low Amplitude (LA)	deg	Mean and SD of minimum angles
	Range of Motion (ROM)	deg	Mean and SD of angular excursion between max and min
	Over-Reaching amplitude (OR)	deg	Mean and SD of peak value exceeding a threshold
Temporal	Peak-to-Peak Time (PPT)	s	Time between adjacent peaks
	Peaks Per Minute (PPM)	Peaks/min	Number of peaks per minute
Velocity based	Raising Velocity (RV)	deg/s	Mean and SD of speed during raising phase
Spectral	Peak Frequency (PF)	Hz	Frequency corresponding to maximum power in the PSD

Table 3.3: Unilateral outcome parameters divided by their relative domain. A brief description and their unit are also reported.

Domain	Parameter	Unit	Description
Angular	High Amplitude Difference (ΔHA)	deg	Difference between HA_{left} and HA_{right}
	Low Amplitude Difference (ΔLA)	deg	Difference between LA_{left} and LA_{right}
	Range of Motion Difference (ΔROM)	deg	Difference between ROM_{left} and ROM_{right}
Temporal	Inter-peak Interval Difference (IPID)	s	Time difference between consecutive peaks of left and right arms
	Peaks Time Lag (PTL)	s	Average time difference between the occurrences of peaks in left and right arms
	Cross-Correlation Lag (CCL)	s	Temporal offset at which the cross-correlation function is maximum
Velocity based	Raising Velocity Difference (ΔRV)	deg/s	Difference between the mean raising velocity of each arm
	Mean Velocity Difference (ΔMV)	deg/s	Difference between the overall mean velocity of each arm

Table 3.4: Bilateral outcome parameters divided by their relative domain. A brief description and their unit are also reported.

3.3.1 Unilateral Outcome Parameters

High Amplitude

The *High Amplitude* (HA), expressed in degrees, represents the average of the peak joint angles reached by the subject across all repetitions of a given movement. This parameter is used to evaluate the subject's ability to adequately raise the upper limb and is particularly informative when compared with the target angle required for the correct task execution.

The value is computed by identifying the maximum peaks using the segmentation algorithm previously described, and calculating their arithmetic mean.

The non-averaged (i.e., single repetition values) values of this parameter is used exclusively for validation purposes.

Low Amplitude

The *Low Amplitude* (LA), also expressed in degrees, represents the average of the minimum joint angles reached by the subject during the execution of the various repetitions of the task. Ideally, this value should be close to zero (or to a possible signal's baseline offset), and it reflects the subject's ability to fully complete each movement cycle without employing compensatory strategies, such as keeping the limb raised between one repetition and the other.

The calculation is done by identifying the local minima with the segmentation algorithm previously described, and subsequently calculating the arithmetic mean.

The non-averaged version of this parameter (i.e., single repetition values) is used only for validation purposes.

Range of Motion

Range of Motion (ROM), expressed in degrees, quantifies the average angular excursion performed by the subject during the task execution. An high ROM value indicates a good motor ability to correctly perform the exercise, while lower values may reflect potential motor limitations.

This parameter is computed as the average of the differences between each maximum peak and its immediately preceding local minimum (the minimum on the left). The use of the left minimum is justified by the fact that this represents the starting point of the movement of the single repetition. Since ROM is calculated as a difference, any offsets present in the signal are effectively removed.

The version without average is used exclusively for validation purposes.

Peak-to-Peak Time

The *Peak-to-Peak Time*, (PPT) expressed in seconds, represents the average time interval between two consecutive angular peaks across all performed repetitions. This parameter provides an indirect measure of movement execution speed: longer intervals may reflect impaired motor performance or reduced coordination.

The value is obtained by averaging the time intervals between each pair of consecutive maxima identified during the task.

The repetition-based version is used only for validation purposes.

Over-Reaching Amplitude

The *Over-Reaching Amplitude* (OR) parameter, measured in degrees, quantifies the average amount by which a subject exceeds a predefined angular threshold during task execution. This threshold was set at 50°, but it can be modified according to specific clinical requirements. This parameter reflects the value of motor impairment compared to the set threshold: higher values suggest better motor condition.

The value is computed as the arithmetic mean of the difference between each angular peak and the predefined threshold.

This parameter is also calculated individually for each repetition during the validation phase.

Peaks Per Minute

The *Peaks Per Minute* (PPM) parameter indicates the frequency of repetitions performed, expressed as the number of peaks detected per minute. It serves as an indirect measure of movement speed and fluidity: higher values are generally associated with better motor performance.

The formula used is the following:

$$\text{PPM} = \frac{N_{\max}}{T} \times 60$$

Here:

- $N_{\max}$ is the total number of maximum peaks detected;
- T is the total duration of the exercise, expressed in seconds.

Raising Velocity

The *Raising Velocity* (RV), represents the average speed of raising the limb during task execution. High values of this parameter indicate a good execution speed,

while low values may reflect motor difficulties. This parameter, expressed in degrees per second ($\frac{\text{deg}}{s}$), is computed as the average of the ratios between the ROM and the required time to reach the angular peak starting from the preceding minimum (i.e., the left minimum). The points used for this calculation are the same as those employed in the ROM definition:

$$\text{RV} = \frac{1}{N} \sum_{i=1}^{N} \frac{\text{ROM}_i}{t_{\text{max},i} - t_{\text{min left},i}}$$

In this formula:

- ROM_i is the Range of Motion of the i-*th* repetition;
- t_{max}, i and t_{min}, i are respectively the time instants of the maximum peak of the i-*th* repetition and its preceding left minimum;

The non-averaged version of this parameter is used for validation purposes.

Peak Frequency

Peak Frequency (PF) represents the frequency corresponding to the maximum power within the spectrum of the angular trajectory. It provides the dominant frequency of the repetitions performed: higher values indicate faster and more regular movements, while lower values could reflect slower executions or possible pauses between repetitions.

To compute this metric, the power spectral density (PSD) of the angular signal is calculated using the Welch method [61], applied on analysis windows of 250 samples. The parameter is expressed in Hz.

3.3.2 Bilateral coordination parameters

These parameters are computed exclusively for bilateral tasks, i.e., simultaneous and alternating raises, since they serve as specific indicators to describe the coordination between upper limbs movements. This aspect of motor perfomance is analyzed through angular, temporal, and velocity-based metrics. These coordination metrics are designed to complement the analysis of the parameters calculated individually for each limb, providing a complete and in-depth assessment of the motor and coordination skills.

High Amplitude Difference

The *High Amplitude Difference* (ΔHA) represents the difference between the mean maximum amplitudes reached by the right and left arms during the execution of a

bilateral task. This parameter, expressed in degrees, reflects the subject's ability to use both limbs symmetrically. Values close to zero suggest balanced motor performance, while higher values may indicate impairment or reduced functionality in one limb compared to the other. ΔHA is calculated as the absolute difference between the mean values of the High Amplitudes of the right and left arms.

The non-averaged version of this parameter is used only for validation purposes.

Low Amplitude Difference

The *Low Amplitude Difference* (ΔLA), also expressed in degrees, represents the difference between the mean minimum values reached by each arm during the execution of a bilateral task. This indicator reflects the subject's ability to control the limbs during the final phase of the repetition. Values close to zero indicate a balanced use of the limbs, whereas high values may suggest a reduced control of one limb compared to the other. This parameter is calculated as the absolute difference between the mean minimum amplitudes of the two limbs ($|LA_{\text{mean,left}} - LA_{\text{mean,right}}|$).

The version relating to each single repetition is used for validation purposes.

ROM Difference

The *ROM Difference* (ΔROM) parameter quantifies the deviation between the mean angular excursions (ROM) of each arm. Values close to zero indicate good symmetry and effective bilateral motor control, while values tending to higher values suggest the presence of one limb that fails to perform the movement as completely as the other.

This parameter is computed as the absolute difference between the mean ROMs of the two limbs.

$$\Delta ROM = |ROM_{\text{mean,left}} - ROM_{\text{mean,right}}|$$

The single repetition version is used for validation purposes.

Inter Peak Interval Difference

The *Inter Peak Interval Difference* (IPID), expressed in seconds, quantifies the time difference between consecutive peaks of the left and right arms during bilateral tasks. It is an indicator of the subject's ability to maintain temporal synchronization between limbs. Values close to 0 indicate that the patient tends to perform repetitions with temporal synchronization between the arms, while higher values suggest a possible motor or coordination deficit. This parameter is calculated as the absolute difference between the average Peak-to-Peak Time of the left arm and the right arms:

$$\text{IPID} = |PPT_{\text{mean,left}} - PPT_{\text{mean,right}}|$$

where $PPT_{\text{mean,left}}$ and $PPT_{\text{mean,right}}$ represent the average Peak-to-Peak Times of the left and right arms, respectively.

The repetition-based version of this parameter is used for validation purposes.

Peak Time Lag

The *Peak Time Lag* (PTL) parameter measures the average time interval between the occurrence of the angular maximum peak of one arm and the corresponding peak of the controlateral arm. This metric, expressed in seconds, reflects the subject's ability to maintain proper inter-arm coordination with respect to the required motor task. In alternating raises, a larger time interval between the peaks is expected, while in simultaneous raises this interval should be minimal and tending to zero. The parameter is computed using the following formula:

$$\text{PTL} = \frac{1}{N}\sum_{i=1}^{N} |t_{\text{max,left},i} - t_{\text{max,right},i}|$$

Where $t_{\text{max,left},i}$ and $t_{\text{max,right},i}$ denote respectively the time instants of the angular maxima for the left and right arms, as identified by the segmentation algorithm.

Also in this case, the version calculated for a single repetition is used for validation purposes.

Cross Correlation Lag

The *Cross Correlation Lag* (CCL), expressed in seconds, defines the temporal offset at which the cross-correlation function, calculated between the angular trajectories of the right and left arm, reaches its maximum value. This parameter provides a global measure of bilateral coordination and does not depend on segmentation, evaluating the overall synchrony between the two angular signals. In simultaneous bilateral tasks, a CCL value close to zero is expected, while in alternating tasks, CCL value should be similar to PTL parameter: deviations from these expectations may indicate difficulties in maintaining the synchrony required by the motor task.

Raising Velocity Difference

The *Raising Velocity Difference* (ΔRV) parameter, expressed in degrees per second ($\frac{\text{deg}}{s}$), represents the disparity in the average speed with which the two arms are raised during the task execution . A value close to zero indicates good symmetry and bilateral motor coordination, while higher values suggest possible motor difficulties

in one of the limbs. It is calculated as the absolute difference of the average raising velocities between the left and right arms:

$$\Delta RV = |RV_{\text{mean,sx}} - RV_{\text{mean,dx}}|$$

The single repetition-based of this parameter is used only for validation purposes.

Mean Velocity Difference

The Mean Velocity Difference (ΔMV), also expressed in degrees per second ($\frac{\text{deg}}{s}$), describes the absolute difference between the overall mean velocities of the two arms calculated across entire duration of the angular trajectory. It is a global measure of the symmetry of the execution velocity. Values close to zero indicate that the subject performs the task with similar velocities for both arms, whereas higher values may denote significant motor asymmetries. This parameter is computed as the absolute difference between the overall mean velocity of the left and right arms.

3.4 Validation protocol

The quality, robustness, and reliability of the data collected through *Palestra*, are key factors in assessing its limitations and potential applications. To investigate these aspects, a comparison with the reference system was conducted, through an agreement analysis between these two systems. For this purpose, five healthy young adults were recruited, whose characteristics are summarized in Table 3.5.

Subject	Gender	Age	Height [m]	Weight [Kg]	Valid trials (Lateral/Frontal)
1	M	25	1.70	68	8/8
2	F	24	1.60	48	8/7
3	M	24	1.74	78	8/8
4	M	27	1.78	80	8/8
5	M	25	1.68	60	8/8

Table 3.5: Participant characteristics and respective valid trials acquired.

Experimental protocol

The reference system (i.e., the gold standard) employed was an *Optitrack* optoelectronic system, composed of six infrared (IR) cameras (1280 x 1224 px resolution) operating at a sampling rate set to 120 Hz. The system covered a working volume close to 6 x 4 x 3 m. To enable motion tracking, a configuration of nine reflective markers (15 mm in diameter) was applied to each subject, emulating the position of the landmarks predicted by Google Media Pipe Pose (see Figure 2.1). Specifically, two markers were placed on the shoulders, elbows, wrists and hips, and one more

marker was positioned on the head. Subjects were required to wear neutral-colored, tight-fitting clothing to minimize motion artifacts. Data acquisition was conducted under controlled lighting conditions. An illustration of the markerization is shown in Figure 3.13.

Figure 3.13: Example of marker placement. The subject is wearing neutral-colored and tight-fitted clothes to minimize motion artifacts.

Validation recordings were performed simultaneously between the two systems, following an acquisition protocol that required the execution of all movements, both lateral and frontal, in every modality (single-arm right and left, simultaneous and alternating raises). Each participant completed two consecutive sessions, separated by a brief rest interval.

The exergame was configured in "controlled" mode, with a set of 10 repetitions per task. A total of 80 trials were thus collected: 40 related to lateral tasks and the remaining 40 to frontal tasks. Only one corrupted frontal trial was discarded (the firts trial of subject 2), due to technical problems during the task execution; therefore, 79 trials were included in the subsequent analysis. The data obtained through the exergame will be referred to as **MP**, while the optoelectronic data will be referred to as **OPT**.

Data pre-processing

To allow an accurate comparison between the two systems, the pipeline described in Section 3.2 was also applied to the OPT signal, with the following three exceptions:

1. **Initial interpolation of trajectories**. During task execution, occasionally the tracking of some markers was lost due to occlusions caused by clothes during movements, resulting in missing values (NaN) within the matrix containing the world trajectories. To obtain complete and comparable signals, a linear interpolation was applied via the MATLAB *fillmissing* function.

2. **Downsampling**. To standardize the sampling frequency between the two systems, the OPT signals were downsampled (from 120 to 50 Hz) using the MATLAB *interp1* functions, as described in the pipeline.

3. **Temporal alignment**. To determine the relative delay between the MP and OPT angular trajectories, the *finddelay* function was used. Once the time shift was computed, the two signals were realigned and subsequently their size was equalized, based on the shortest one. The delay correction was essential for computing error metrics and for the Bland-Altman analysis performed on the entire trajectories.

Figure 3.14 illustrates examples of pre-processed and overlaid angular trajectories are shown.

As shown in this figure, temporal alignment is nearly perfect, in both lateral and frontal tasks; whereas spatial alignment reveals greater variability. In particular, better spatial alignment is observed in lateral tasks, while frontal tasks tend to show more variation, due to tracking difficulties and random fluctuations in the signals. In some cases, the spatial alignment is highly accurate, i.e., in the first plot of the second raw, but in some cases greater variability is evident. This variability may depend on limitations in Google MediaPipe Pose tracking, potentially influenced by factors such as the subject's clothing, execution speed and anthropometrich measures.

Finally, the same outcome parameters described above were extracted for the two systems to create the datasets necessary for statistical analysis.

These parameters were calculated pursuing two different strategies, aimed at evaluating the also robustness of the segmentation algorithm:

1. Applying the segmentation algorithm independently to both signals.

2. Using the segmentation derived from the OPT signal (i.e., the positions of its maximum and minimum points detected) on the aligned MP signal.

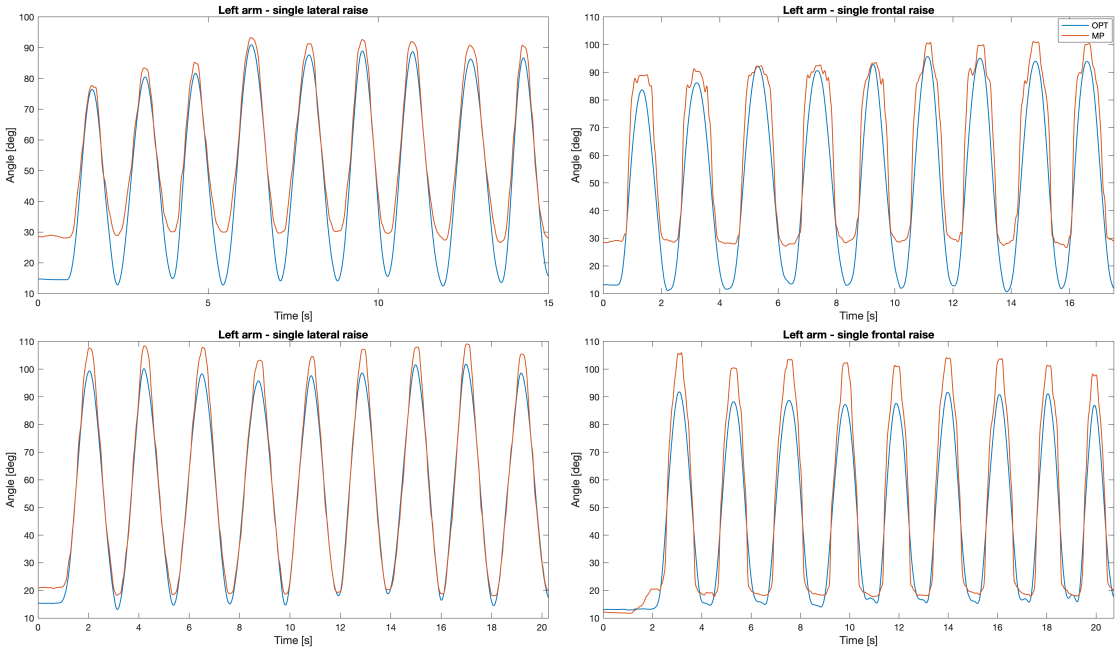

Figure 3.14: Aligned angular trajectories obtained during single left-arm raises, both lateral and frontal. The first raw shows signals from Trial 1 of subject 1, while the second raw illustrates trajectories from Trial 1 of subject 5.

3.4.1 Statistical Analysis and dataset creation

To assess the agreement between the two systems, a series of statistical analysis were conducted, both on the full trajectories and on the extracted parameters. The analyses were performed using Jamovi [62] and MATLAB.

Overall trajectories analysis

To conduct a robust concordance study on the entire trajectories, the initial static phase of the signal (i.e., the first three seconds of the exergame recordings) was removed, to avoid introducing bias into the computed parameters. The agreement between the systems was evaluated with the following metrics:

- **Root Mean Square Error (RMSE)**. This metric quantifies the mean squared deviation between two measures, serving as an indicator of the overall discrepancy.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\theta_i^{\mathrm{OPT}} - \theta_i^{\mathrm{MP}}\right)^2}$$

  θ_i^{OPT} and θ_i^{MP} represent the angular values obtained from the optoelectronic

system and with the exergame, respectively, at the i-*th* time point. This metric is expressed in degrees.

- **Percentual Root Mean Square Error (PRMSE)**. This metric is the normalized version of the RMSE, expressed as percentage relative to the reference signal (OPT).

$$\text{PRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\theta_i^{\text{OPT}} - \theta_i^{\text{MP}}}{\theta_i^{\text{OPT}}}\right)^2} \cdot 100$$

Analysis on extracted outcome parameters

For parameters extracted as mean values, as well as those derived on the entire trajectory (e.g., the spectral parameters), the **relative percentage difference**, **RMSE** and **PRMSE** between the mean OPT and MP values were calculated. The aim of this analysis is to provide an initial and global overview of the final system's performance.

For each extracted parameter, descriptive statistics were performed, including the **mean**, **standard deviation**, **median**, **minimum**, **maximum values**, **sample size** and the **p-value** for testing the normality of the distribution. To assess the concordance between the two measures, **Bland-Altman** plots were generated. Subsequently, either the **Pearson or Spearman correlation coefficient** was computed (depending on the data distribution), to analyse the correlation between the two set of measurements.

The Bland-Altman analysis [63] facilitates both graphical and numerical comparison between two measurement systems. The difference between paired observations is displayed on the Y axis, in relation to their mean, which is shown on the X axis. The plot includes three reference lines: the central line identifies the *bias*, defined as the mean difference between the two systems, while the upper and lower lines are the *Limits of Agreement* (LoA), indicating the interval within which the 95% of the difference must be contained.

Dataset organization

Three main datasets were constructed for each performed analysis: one for the mean parameter analysis, one for parameters computed with the standard segmentation and one for parameters extracted using OPT-based segmentation. Each dataset was further subdivided in function of the execution type (lateral and frontal) and by the type of parameter analyzed (unilateral and bilateral parameters). Table 3.6 summarize the dataset division, reporting the number of samples in each subset. Each subset is representative of a specific parameter, for example, one subset for HA another for ROM, etc. As it can be observed, dataset sizes are not

uniform, especially those related to frontal tasks, where certain repetitions were not correctly detected. Some parameters, such as the PPT (time difference between adjacent peaks) have a lower dimensionality than subset size, due to their analytical definition. These dimensionalities are reported directly in *Results* chapter.

Dataset division	**Mean parameters**		**Standard segmentation**		**OPT-based segmentation**	
Task type	Lateral	Frontal	Lateral	Frontal	Lateral	Frontal
Unilateral	60	59	440	420	440	420
Bilateral	20	20	130	129	130	129

Table 3.6: Overview of Dataset division. The subset size is reported, divided by analysis type (Mean parameters, standard segmentation, and OPT-based segmentation) by execution type (Lateral and frontal raises) and by parameter type (unilateral and bilateral).

3.5 Parkinsonian participants analysis

To conduct an initial real-world evaluation, data were collected from 12 parkinsonian subjects, who present varying levels of impairment. Participants were recruited at a local association of parkinsonian subjects involved in ELEVATOR project, where they performed all the tasks under the supervision of a neurologist. In this experimental test, exergame was configured in "controlled" mode, setting a limit of five repetitions per task. Subjects were instructed to complete all the proposed tasks, thus acquiring a total of 96 trials. However, 2 trials related to the alternating lateral raises were excluded as the subjects were unable to complete the task. In conclusion, the eligible signals are 94: 46 corresponding to lateral raises and 48 to frontal raises.

All eligible trials were analyzed using the pipeline described in section 3.2, to extract the outcome parameters reported in Tables 3.3 and 3.4. The trials were analyzed separately by movement type (lateral and frontal raises) and task modality (single-arm, alternating, and simultaneous raises).

Figures 3.15 shown angular trajectories obtained from a parkinsonian subject.

As can be observed, the trajectories maintain a great signal quality, even in frontal raises. This may be attributed to the slower execution of the movement compared to healthy subjects.

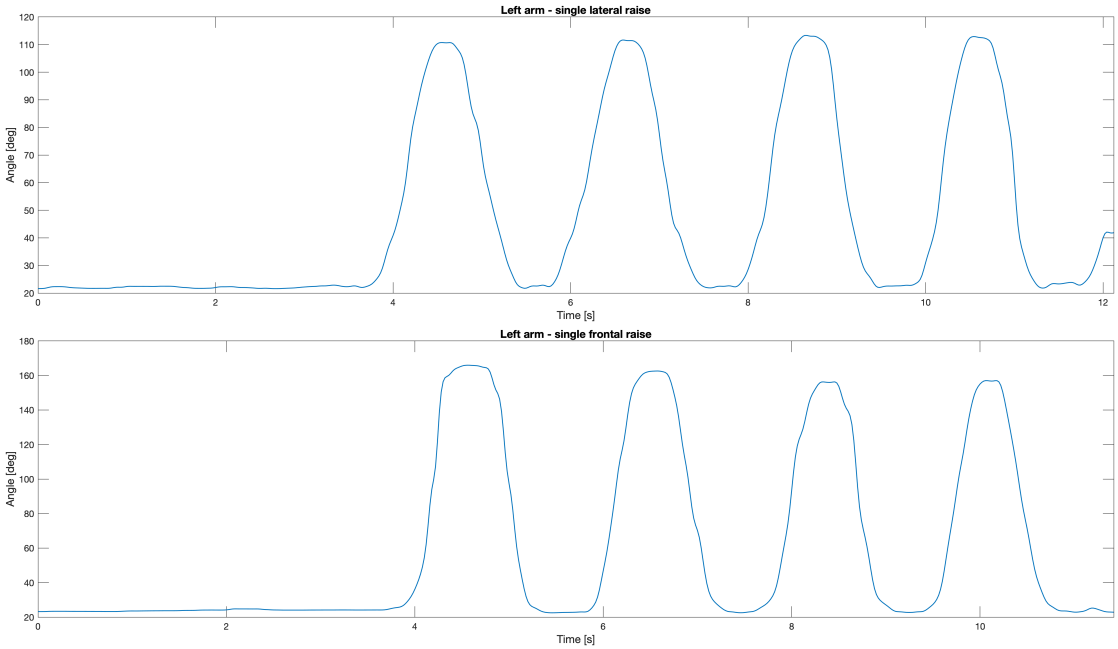

Figure 3.15: Example of angular trajectories for single left arm lateral (top) and frontal (bottom) raises, acquired from a parkinsonian partecipant.

Chapter 4

Results

This section presents the results of the validation analysis, followed by a descriptive overview of the parameters extracted from the parkinsonian subjects.

4.1 Overall 3D trajectories analysis

In this section, the trajectories obtained in each trial using the exergame and the MediaPipe (MP) framework are compared with the corresponding trajectories simultaneously acquired by the optoelectronic system (OPT). The comparison is based on standard error metrics such as RMSE and PRMSE, in addition to the evaluation of linear correlation using the Pearson coefficient.

The analysis is structured to provide insights at increasing level of detail. Initially, all trajectories are considered, without distinguishing between movement directions or task types. Subsequently, the analysis considered separately lateral and frontal raises, and finally each individual task performed. This choice was employed to highlight different behaviours of the MP tracking model according different movement conditions

The first result presented regards the comparison between all the 119 3D trajectories captured by both MP and OPT, during the exergame executions.

N	**RMSE [deg]**	**PRMSE [%]**	$\boldsymbol{\rho}$
119	9.040 ± 3.706	41.256 ± 24.596	0.988 ± 0.016

Table 4.1: Error metrics and Pearson correlation coefficient ($mean \pm SD$) evaluated over all trajectories. N indicates the number of trials considered.

Table 4.1 presents the results of the statistical analysis performed on the angular trajectories. From a first general comparison, the average value of the RMSE computed across all the trajectories (both lateral and frontal) is $9.040 \pm 3.706°$.

Considering that the average amplitude of the movements performed during the tasks typically exceeds 90°, this error corresponds to approximately 10% of the mean motion range.

The PRMSE value highlights higher percentage errors, accompanied by a significant standard deviation, primarily due to the differences observed in the minimum points of the angular trajectories. As illustrated in Figure 4.1, in some cases the difference between the minimum values of the OPT and MP signals even exceeds 15°. For instance, if at the i-*th* time instant the OPT signal measure 12° and the MP signal 30°, the PRMSE at that point, computed applying the formula previously described, would be 150%, due to the small denominator (the OPT_i value). By contrast, at the maximum points, the denominator is larger, resulting in lower instantaneous PRMSE values.

In general, the RMSE values suggest that the exergame provides reasonably accurate estimates, also considering the intended application where a very high precision is not required. However, the PRMSE indicates that caution is needed at the minimum point of the signals. To conclude, the linear correlation coefficient is close to 1, demonstrating that the trajectories tend to follow the same overall trend.

Going into detail, it is possible to divide the trajectories into lateral and frontal ones, computing the same error metrics separately.

Movement type	**N**	**RMSE [deg]**	**PRMSE [%]**	$\boldsymbol{\rho}$
Lateral	60	7.293 ± 3.423	36.658 ± 23.033	0.990 ± 0.007
Frontal	59	10.816 ± 3.107	45.932 ± 25.436	0.971 ± 0.015

Table 4.2: Error metrics and Pearson correlation coefficient ($mean \pm SD$) evaluated for each movement type. N indicates the number of trials considered

In relation to lateral raises, the RMSE values are lower compared to those observed in frontal ones, and a similar trend is also noticeable in the PRMSE. The Pearson correlation coefficient, although high in both cases, exhibits greater values in the lateral tasks. These findings confirm that lateral movements are more accurate than frontal ones.

The boxplots (Figure 4.2) further support the numerical results: lateral angular trajectories are, on average, more accurate and present lower variability against frontal ones, for both RMSE and PRMSE. The presence of outliers is evident in the graphs: for RMSE these may be attributed to localized signal errors, while for the PRMSE they can be explained by the denominator-effect previously discussed.

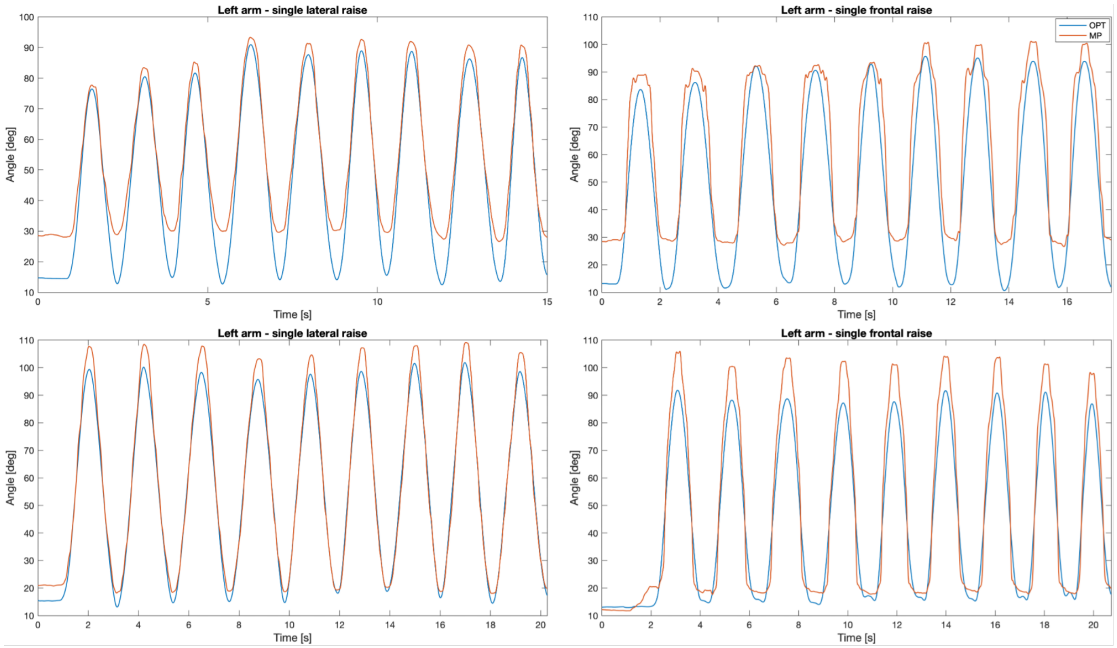

Figure 4.1: Overlapped single-arm raise trajectories. The OPT signals (blue) represent the reference system, while MP signals (red) are the estimated trajectories. The most pronounced discrepancies are observed at the minima and maxima of angular trajectories. In the first raw are reported an example of alignment for both lateral and frontal raises, where the minimum values are relatively distant between the two signals. In the second raw, the alignment is nearly perfect, with minimum values notably overlapped and maximum values closely aligned.

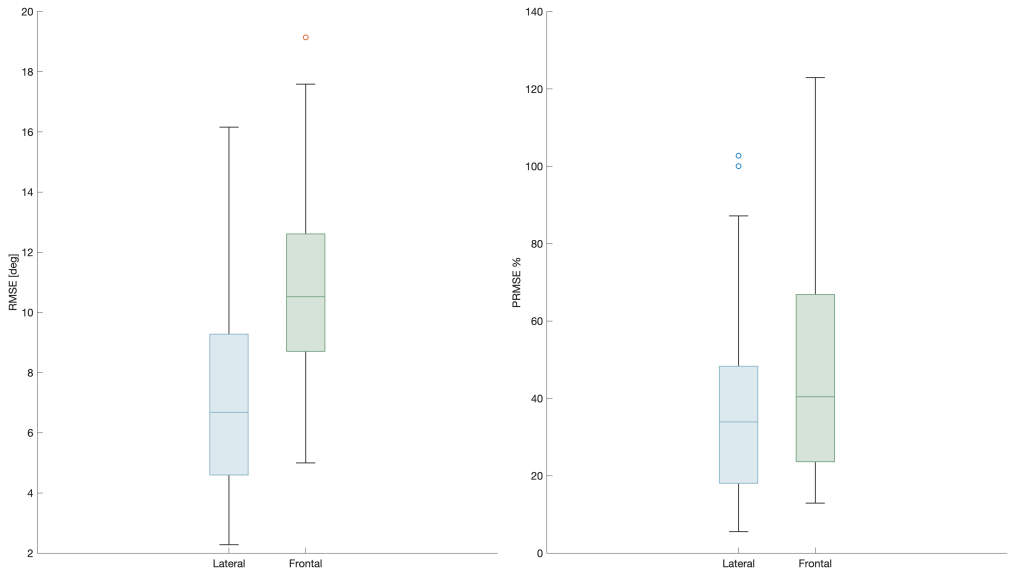

Figure 4.2: Boxplots of RMSE (left) and PRMSE (right) for lateral and frontal movements. Both metrics evidence greater accuracy in lateral movement, as highlighted by lower medians and variability.

Continuing with the analysis, the trajectories were further subdivided by task type, to identify whether some exercises are more challenging to track and if this results in less accurate trajectories.

Task (lateral)	N	RMSE [deg]	PRMSE [%]	ρ
Single-arm	20	6.159 ± 2.767	27.769 ± 16.134	0.997 ± 0.004
Alternating	20	8.504 ± 3.711	52.751 ± 27.119	0.988 ± 0.007
Simultaneous	20	7.215 ± 3.472	29.455 ± 15.535	0.999 ± 0.003

Table 4.3: Error metrics and Pearson correlation coefficient for each lateral task ($mean \pm SD$). N indicates the number of trials considered.

Task (frontal)	N	RMSE [deg]	PRMSE [%]	ρ
Single-arm	19	10.145 ± 3.458	41.223 ± 25.260	0.973 ± 0.014
Alternating	20	9.755 ± 2.434	50.224 ± 28.779	0.972 ± 0.015
Simultaneous	20	12.516 ± 2.747	46.112 ± 22.375	0.969 ± 0.017

Table 4.4: Error metrics and Pearson correlation coefficient for each frontal task ($mean \pm SD$). N indicates the number of trials considered.

Table 4.3 and Table 4.4 report the mean values and corresponding standard deviations of the computed metrics, for all types of tasks performed, respectively grouped as lateral and frontal raises.

In general, all error metrics are lower in lateral tasks compared to frontal ones, confirming a better tracking performance in lateral movement. A marked difference is observed in simultaneous raises, where the mean RMSE is $7.215 \pm 3.472°$ for lateral movements, and raises to $12.516 \pm 2.747°$ in frontal ones.

Analysing the individual tasks, lateral raises exhibit the lowest RMSE in single-arm movements ($6.159 \pm 2.767°$), followed by simultaneous ($7.215 \pm 3.472°$), and alternating raises ($8.504 \pm 3.711°$). The Pearson correlation coefficient remains steadily high, close to 1, with a highest value observed in the simultaneous raises (0.999 ± 0.003).

In contrast, in frontal raises, the RMSE is lower in alternating raises ($9.755 \pm 2.434°$) followed by single-arm ($10.145 \pm 3.458°$) and simultaneous raises ($12.516 \pm 2.747°$). This finding seems to be in contrast with the pattern observed in lateral movements; however, it can be explained by the oscillatory transients characteristic of the alternating raise. As described in the *Methods* section, alternating frontal raises are characterized by a concave oscillation: this leads to a smaller deviation from the OPT signal, and consequently, a decrease in error values within the transitory interval. In contrast, alternating lateral raises display a convex undulatory pattern, which increases the mismatch between the two signals. This phenomenon

is illustrated in Figure 4.4.

The PRMSE, however, is not affected by this phenomenon, as it is high in both lateral ($52.751 \pm 27.119\,\%$) and frontal ($50.224 \pm 28.779\,\%$) alternating tasks, indicating that percentage errors are more marked at the signal minima, where transient effects typically occur.

The Pearson correlation coefficient in the simultaneous frontal raises (0.969 ± 0.017) shows the smallest value among all tasks, although it maintains a value close to 1 (i.e., strong correlation with OPT signals).

Figure 4.3 displays the distribution of error metrics across tasks, comparing lateral and frontal executions. Lateral raises record lower values for all metrics in all task types, except for PRMSE in the alternating tasks, which is similar to frontal execution.

In conclusion, these findings confirm that the tracking accuracy is partly influenced by both the type of motor task and by the plane in which the movement is performed:

- Movements in the lateral direction are more accurately acquired by the MP tracking model compared to those captured in the frontal direction;
- Single-arm movements are the better tracked in the lateral raises;
- Alternating movements presents great stability in frontal execution.

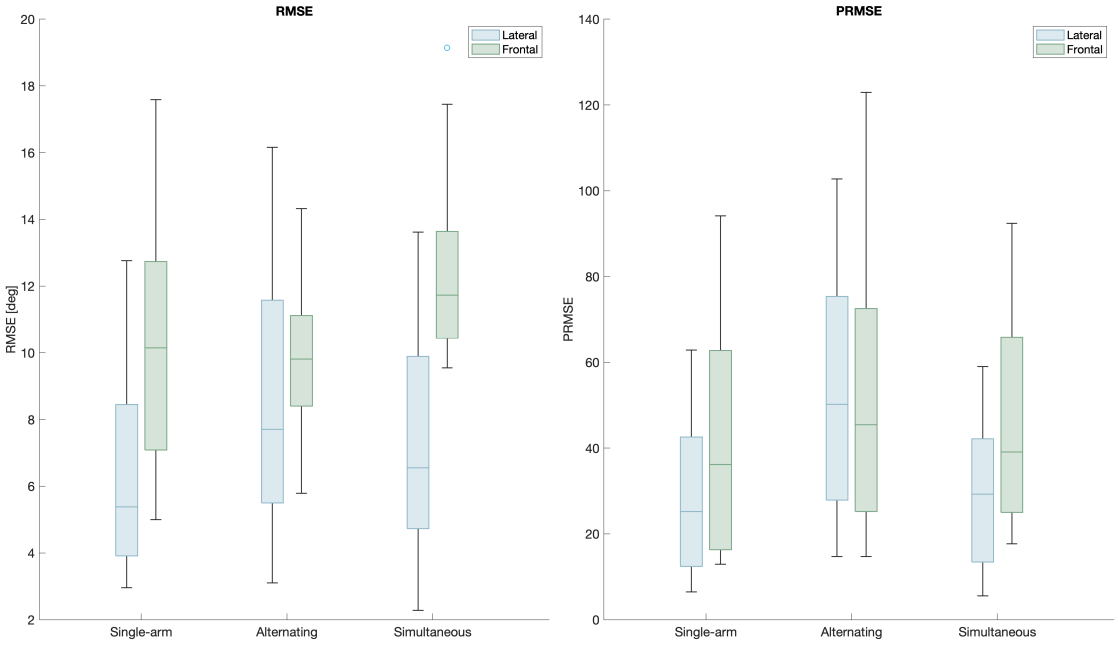

Figure 4.3: Boxplots of RMSE (left) and PRMSE (rigth) for each task. Both metrics demonstrate greater accuracy in lateral tasks, except for the PRMSE of the alternating frontal raises.

4.2 Analisys on mean outcome parameters

After the analysis of overall trajectories, the validation analysis was performed on kinematic parameters extracted from 3D trajectories, considering two levels of detail:

- First, the averaged parameters, in addition to those derived from the entire signal (e.g., spectral parameters) were evaluated. These outcomes represent the aggregate data that would be available in a real-world application of the exergame.
- Second, a more specific analysis of the same parameters estimated for each movement repetition was performed, to take advantage of the larger dataset obtained through this approach.

This preliminary analysis provides an initial estimate of the system's overall performance. Based on the results obtained from the trajectory analysis, the aggregate (i.e., the averaged) parameters extracted were divided into lateral and frontal groups. Bilateral parameters were similarly grouped.

For each parameter, the mean and standard deviation (SD) were computed separately for both OPT and MP signals. In addition, the RMSE, the PRMSE and the relative percentage difference between the two measurements were calculated.

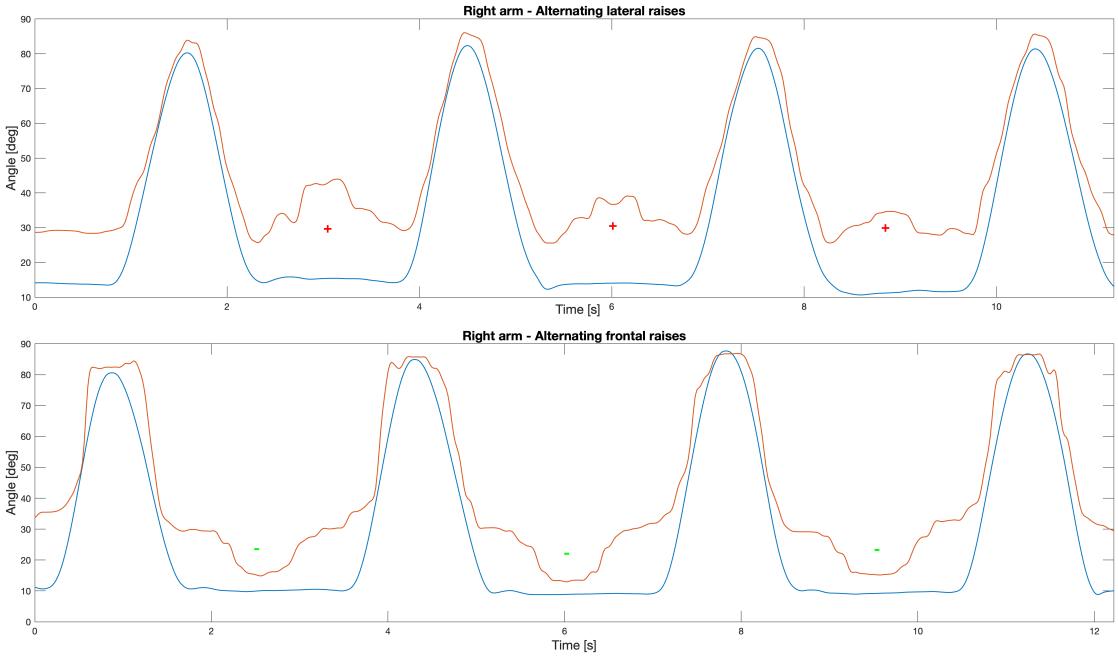

Figure 4.4: Example of transitory phenomena in the alternating task. During alternating lateral raises (top) the undulation is convex (+) in quiet periods, whereas in the frontal execution (bottom) is concave (-).

Unilateral parameters analysis

In this section the mean unilateral parameters (i.e., those computed from single-arm executions, across all task types) are reported. The first table presents the values related to lateral raises, while the second those for frontal raises.

Domain	Parameter	N	MP	OPT	RMSE	PRMSE	Δ%
Angular	HA [deg]	60	92.80 ± 6.41	88.73 ± 7.01	6.03	7.01	−4.59
	LA [deg]	60	22.26 ± 3.56	14.78 ± 2.82	9.26	72.14	−50.57
	ROM [deg]	60	70.50 ± 7.43	73.81 ± 5.49	5.10	7.06	4.49
	OR [deg]	60	42.80 ± 6.41	38.73 ± 7.01	6.03	17.14	−10.52
Temporal	PPT [s]	60	2.65 ± 0.90	2.65 ± 0.90	0.02	0.49	−0.03
	PPM [peaks/min]	60	25.38 ± 6.60	25.38 ± 6.60	0.00	0.00	0.00
Velocity-based	RV [deg/s]	60	67.37 ± 11.21	73.45 ± 6.78	8.78	12.77	8.28
Spectral	PF [Hz]	60	0.40 ± 0.13	0.40 ± 0.14	0.04	18.26	−1.68

Table 4.5: Mean values (±SD) and error metrics for each outcome parameter, computed with both MP and OPT, for unilateral **lateral** raises. Negative Δ% values reveal that MP overestimates the parameter compared to OPT. N indicates the number of parameters considered.

Table 4.5 and Table 4.6 report the mean values of the parameters extracted from both systems, along with the RMSE, PRMSE and the relative percentage difference, making a global comparison possible.

Domain	Parameter	N	MP	OPT	RMSE	PRMSE	Δ%
Angular	HA [deg]	59	95.67 ± 10.08	83.61 ± 7.30	14.43	17.68	−14.43
	LA [deg]	59	22.07 ± 3.84	13.06 ± 3.31	10.66	102.82	−69.01
	ROM [deg]	59	73.59 ± 10.57	70.95 ± 6.41	10.15	14.29	−3.72
	OR [deg]	59	45.67 ± 10.08	33.61 ± 7.30	14.43	47.48	−35.90
Temporal	PPT [s]	59	2.83 ± 1.00	2.83 ± 1.00	0.02	0.84	0.22
	PPM [peaks/min]	59	24.41 ± 6.44	24.41 ± 6.44	0.00	0.00	0.00
Velocity-based	RV [deg/s]	59	67.93 ± 13.20	69.99 ± 6.47	11.76	16.43	2.94
Spectral	PF [Hz]	59	0.39 ± 0.14	0.39 ± 0.14	0.00	0.00	0.00

Table 4.6: Mean values (±SD) and error metrics for each outcome parameter, computed with both MP and OPT, for unilateral **frontal** raises. Negative Δ% values reveal that MP overestimates the parameter compared to OPT. N indicates the number of parameters considered.

In general, angular and velocity-based parameters exhibit higher RMSE and PRMSE values than temporal parameters, in both lateral and frontal raises.

Specifically, the High Amplitude (HA) parameter shows an RMSE equal to 6.03° in lateral raises, increasing to 14.43° in frontal raises. These values suggest that the trajectories extracted through the exergame in the frontal tasks tend to overestimate the movement amplitude than in lateral ones. This overestimation is evident in both cases, as indicated by the relative percentage difference: -4.59% for lateral raises and -14.43% for frontal raises. The PRMSE follows the same trend: it is higher in frontal movements (17.68%) compared to lateral ones (7.01%).

A similar conclusion can be derived from the Low Amplitude (LA) parameter: the RMSE is 9.26° for lateral raises and 10.66° for frontal raises. The PRMSE is notably high in both cases (72.14% in the lateral movements and 102.82% in the frontals), indicating the difficulty of the MP model in accurately estimating minimum values. This may be partially attributed to the variability introduced by the segmentation algorithm. Furthermore, the percentage difference (Δ%) confirms an overestimation by the MP model in both movements (-50.57% in the lateral raises and -69.01% in the frontal raises).

The Range of Motion (ROM) shows higher precision in lateral raises (RMSE: 5.10°, PRMSE: 7.06%) compared to frontal raises (RMSE: 10.15, PRMSE: 14.29%). In lateral tasks, MP tends to underestimate (Δ%: 4.49) the ROM compared to OPT, although in frontal tasks, a slight overestimation (Δ%: -3.72) is observed. This discrepancy can be attributed to the greater variability of MP trajectories during frontal raises, as suggested by a high standard deviation (±10.57°). This variability is also largely due to the segmentation algorithm, which is less accurate in identifying local minima in signals characterized by less regular morphology (i.e., the trajectories related to frontal movements). This group of signals also includes alternating raises, which are affected by the transient undulation artifact, which certainly contributes to the observed deviations in the aggregate parameter computation.

The Overreaching (OR) parameter follows a similar trend to the HA, as it differs

from the second only by a constant.

However, temporal parameters show high accuracy in both tasks. In particular, the Peak-per-Minute (PPM) parameter presents error metrics equal to zero, while the Peak-to-Peak Time (PPT) yields a RMSE of 0.02 s and a PRMSE of 0.49% in lateral raises, with slightly higher values in frontal raises (PRMSE: 0.84%). The relative percentage difference is below 1% in both cases, confirming the reliability of the MP model in evaluating temporal aspects of the movement.

The Rising Velocity (RV) parameter is more precise in the lateral raises (RMSE: 8.78° and PRMSE: 12.77%), compared to the frontal ones (RMSE: 11.76° and PRMSE: 16.43%), with a greater underestimation observed in the lateral movements (Δ%:8.28) compared to the frontal ones (Δ%: 2.94).

Concerning spectral parameter, slightly better accuracy is achieved in frontal raises, where the RMSE, PRMSE and the Δ% are all equal to 0. Instead, in lateral raises, slightly higher values are reported: RMSE= 0.04 Hz, PRMSE= 18.26 % and Δ%= -1.68.

To confirm this trend, the boxplots, divided between lateral and frontal raises, have been reported. Only parameters characterized by a high value of error metrics have been considered.

In lateral raises (Figure 4.5), the parameters extracted from the MP signals exhibit greater variability compared to those derived from the OPT signals, with the exception of the HA parameter. This increased variability may be attributed to the heterogeneity of the subjects' movements, which the OPT system is capable to detect. Even the outliers present in the OPT distributions, such as the one observed in the the ROM parameter, may be a consequence of this cause.

Figure 4.6 displays the boxplot related to the frontal raises. Here, the variability of the MP-derived parameters is more evident than in OPT data, further confirming the lower reliability of the system in tracking frontal movements and that could impact on the parameter associated with it.

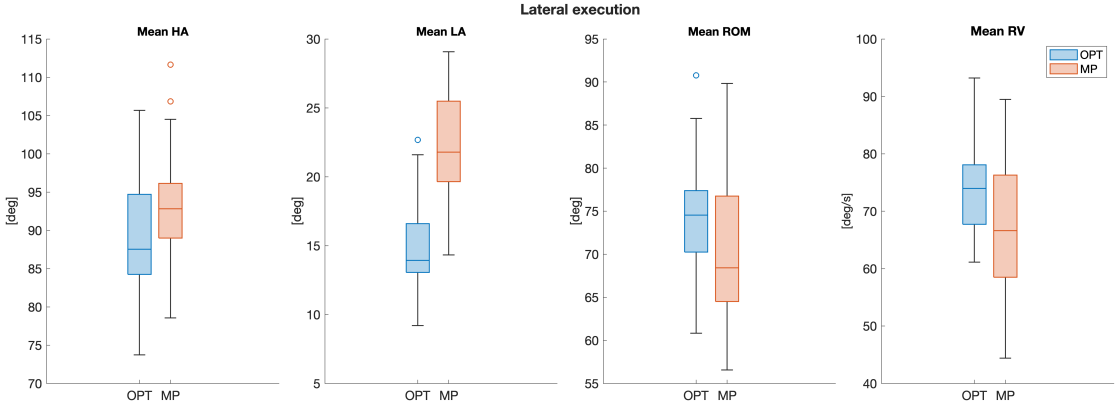

Figure 4.5: Boxplots of outcome parameters with high RMSE and PRMSE values extracted from **lateral** raises. OPT measurements are shown in blue, while MP measurements are displayed in red. Overall, MP presents greater variability compared to the OPT distribution.

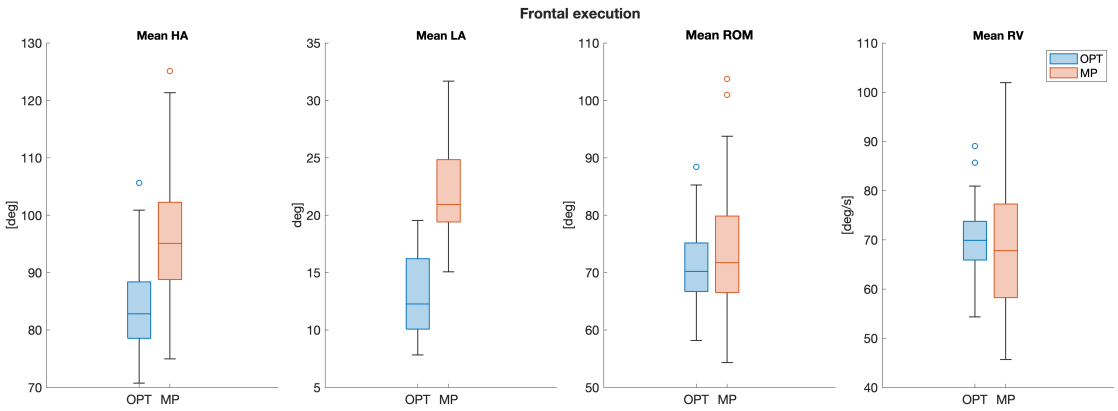

Figure 4.6: Boxplots of outcome parameters with high RMSE and PRMSE values extracted from **frontal** raises. OPT measurements are shown in blue, while MP measurements are displayed in red. Overall, MP presents greater variability compared to the OPT distribution.

Bilateral parameters analysis

In this section the mean bilateral parameters (i.e., those computed as the difference between parameters of right and left arms, across simultaneous and alternating tasks) are reported. The first table presents the values related to lateral raises, while the second those for frontal raises.

Only RMSE and $\Delta\%$ are considered in this analysis, as most parameters are computed as differences, which results in high and non meaningful PRMSE values.

Domain	Parameter	N	MP	OPT	RMSE	Δ%
Angular	ΔHA [deg]	20	5.02 ± 3.51	5.38 ± 3.35	4.03	6.62
	ΔLA [deg]	20	2.63 ± 1.38	2.18 ± 1.31	1.55	−20.35
	ΔROM [deg]	20	4.45 ± 3.54	4.50 ± 2.83	3.28	1.06
Temporal	IPID [s]	20	0.03 ± 0.05	0.03 ± 0.05	0.02	−1.49
	PTL [s]	20	0.96 ± 0.98	0.96 ± 0.99	0.02	−0.52
	CCL [s]	20	0.95 ± 1.00	0.95 ± 1.00	0.01	0.00
Velocity-based	ΔRV [deg/s]	20	4.32 ± 3.03	4.05 ± 3.38	3.22	−6.51
	ΔMV [deg/s]	20	0.08 ± 0.06	0.14 ± 0.15	0.13	37.87

Table 4.7: Mean values (±SD) and error metrics for each outcome parameter, computed with both MP and OPT, for bilateral **lateral** raises. Negative $\Delta\%$ values reveal that MP overestimates the parameter compared to OPT. N indicates the number of parameters considered.

Domain	Parameter	N	MP	OPT	RMSE	Δ%
Angular	ΔHA [deg]	20	9.27 ± 5.14	2.80 ± 2.21	7.61	−231.11
	ΔLA [deg]	20	2.53 ± 1.45	2.07 ± 1.12	1.55	−22.07
	ΔROM [deg]	20	8.16 ± 5.74	2.69 ± 1.86	7.63	−203.97
Temporal	IPID [s]	20	0.09 ± 0.21	0.07 ± 0.20	0.03	−17.69
	PTL [s]	20	0.98 ± 0.96	0.96 ± 0.99	0.07	−2.03
	CCL [s]	20	0.98 ± 1.04	0.98 ± 1.04	0.01	−0.10
Velocity-based	ΔRV [deg/s]	20	10.35 ± 8.37	2.80 ± 2.37	10.53	−269.45
	ΔMV [deg/s]	20	0.25 ± 0.33	0.31 ± 0.56	0.43	21.41

Table 4.8: Mean values (±SD) and error metrics for each outcome parameter, computed with both MP and OPT, for bilateral **frontal** raises. Negative $\Delta\%$ values reveal that MP overestimates the parameter compared to OPT. N indicates the number of parameters considered.

Table 4.7 and Table 4.8 show the average values of the parameters extracted with the two systems, in addition to RMSE and the relative percentage difference, to provide a global comparison.

In general, the discrepancies between the systems are more marked in the frontal tasks than in the lateral ones, due to the higher variability measured in the parameters extracted from the signals.

The ΔHA parameter is characterized by an RMSE of 4.03° in lateral raises and 7.61° in frontal raises. The reduced accuracy of this parameter can be attributed to the heterogeneity of the movements performed during the trial, confirmed by the variability in the averaged OPT values (5.38±3.35° for lateral raises and 2.80±2.21° for frontal raises). MP slightly underestimates the parameter in lateral movements

(Δ%: 6.62), by contrast, in frontal movements there is a relevant overestimation (Δ%: -231.11).

For the ΔLA parameter, the accuracy is comparable between the two systems (RMSE: 1.55°), with a minor overestimation in both movement types (Δ%: -20.35 in lateral tasks and Δ%: -22.07 in frontal tasks).

The ΔROM parameter is characterized by a RMSE of 3.28° in lateral raises and 7.63° in frontal raises. It is notable that in lateral raises MP slightly underestimates the parameter (Δ%: 1.06) whereas in frontal raises a marked overestimation is evident (Δ%: -203.97).

The temporal parameters are on average very accurate in both systems, characterized by an RMSE on the order of few hundredths of a second. For example, the CCL parameter, exhibits an RMSE of 0.01 s in both lateral and frontal tasks, with a slight overestimation in frontal raises (Δ%: -0.10).

The velocity-based parameters reveal more evident discrepancies between the two systems, for ΔRV (RMSE: 3.22 °/s in lateral raises, 10.53 °/s in frontal raises). In the frontal movements, the MP system overestimates compared to OPT (Δ%: -269.45), due to the lower quality of the MP signal, as attested by the high variability recorded (4.32±3.03 °/s in the lateral task, while 10.35±8.37 °/s in the frontal one).

Likewise, in this case, the parameters derived from the lateral exercises demonstrate more robustness than those obtained from the frontal raises. This observation is supported by the boxplots in Figures 4.7 and 4.8, which highlighted a greater variability in the parameters derived from MP compared to OPT. In addition, when the distribution of the parameters extracted with OPT presents a considerable variability comparable to that of MP, it suggest a high heterogeneity in the participants' movements.

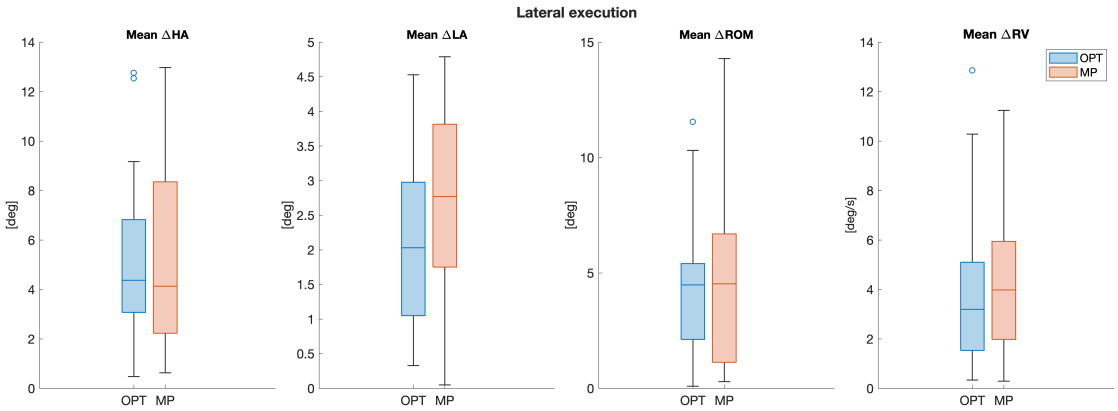

Figure 4.7: Boxplots of bilateral outcome parameters with high RMSE values extracted from **lateral** raises. OPT measurements are shown in blue, while MP measurements are displayed in red. Overall, MP presents greater variability compared to the OPT distribution.

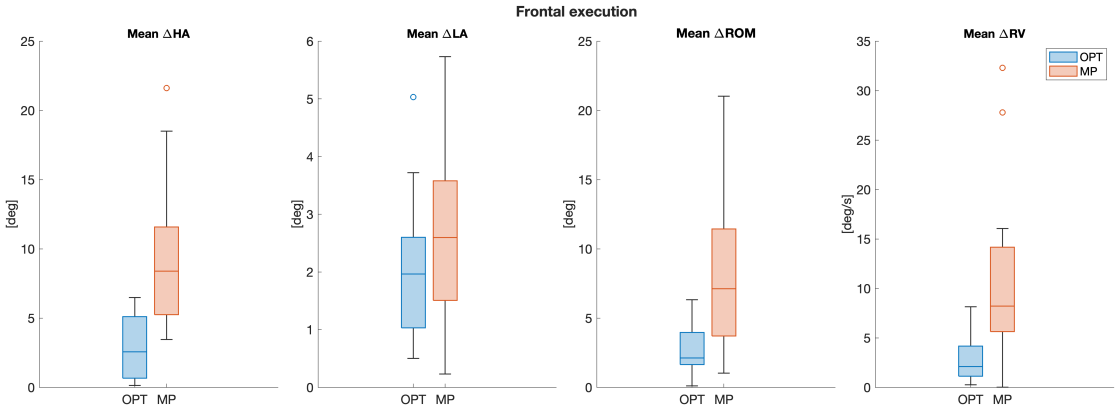

Figure 4.8: Boxplots of bilateral outcome parameters with high RMSE values extracted from **frontal** raises. OPT measurements are shown in blue, while MP measurements are displayed in red. Overall, MP presents greater variability compared to the OPT distribution.

4.3 Analysis on single repetition parameters

To complement the validation analysis, in addition to evaluating average parameters, it was decided to analyse the parameters extracted from each single repetition. This approach increases the statistical power of the study thanks to the larger dataset and enables the application of Bland-Altman analysis. The latter allows the evaluation the agreement between the two measurement systems and for identifying any systematic bias that may be not evident from aggregate parameter analysis.

In addition, the parameters computed for each single repetition allows the construction of dataset in which parameters are calculated using segmentation based on the OPT signal (position of minima and maxima). This serves to assess the robustness of the segmentation algorithm used.

By the findings of the previous analyses, the dataset was divided into two main groups, corresponding to lateral and frontal raises. Each group, in turn, was subdivided into unilateral and bilateral parameters.

The analysis is organized as follows:

1. **Descriptive statistics** for each parameter to characterize its distribution. In addition, the Shapiro-Wilk (SW) test was conducted to assess the normality of the distribution.

2. **Spearman correlation analysis**, to understand the relationship between the two measures.

3. **Bland-Altman analysis** on the most significant parameters to identify the limits of agreement, detect the presence of any systematic bias, and evaluate whether specific tasks fall within or outside these limits.

Finally, the results derived from different segmentation approaches were compared to identify any critical limitations of the proposed segmentation algorithm.

4.3.1 Standard segmentation

This section presents the results obtained from parameters extracted using the standard segmentation method, as described in section 3.2. The analysis is divided between unilateral and bilateral parameters.

Unilateral parameter analysis

The tables below report the descriptive statistics related to the unilateral parameters, which include the mean, standard deviation, median, interquartile range (IQR), maximum, minimum values and the significance of the normality test.

Lateral OPT

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
HA [deg]	440	88.6	7.47	88.5	11.2	109	65.1	0.026
LA [deg]	440	14.8	3.02	14.2	4.13	23.6	8.36	<.001
ROM [deg]	440	73.7	6.12	74.2	7.59	97.0	54.1	0.002
RV [deg/s]	440	73.4	7.29	73.4	11.2	101	54.9	0.01
PPT [s]	380	2.36	0.720	2.20	0.605	4.84	1.46	<.001
OR [deg]	440	38.6	7.47	38.5	11.2	59.00	15.1	0.026

Table 4.9: Descriptive statistic of OPT values for lateral execution. P-values less than 0.05 indicate a non-normal distribution. N indicates the number of parameters considered.

Lateral MP

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
HA [deg]	440	94.0	7.25	93.2	10.5	114	70.6	0.003
LA [deg]	440	22.3	3.47	21.8	5.45	30.5	13.1	<.001
ROM [deg]	440	71.7	8.44	70.9	13.1	91.5	49.6	<.001
RV [deg/s]	440	68.9	12.5	69.4	20.1	97.9	39.7	<.001
PPT [s]	380	2.36	0.721	2.20	0.605	4.88	1.46	<.001
OR [deg]	440	44.0	7.25	43.2	10.5	64.4	20.6	0.003

Table 4.10: Descriptive statistic of MP values for lateral execution. P-values less than 0.05 indicate a non-normal distribution. N indicates the number of parameters considered.

Frontal OPT

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
HA [deg]	420	83.4	7.45	83.0	9.86	110	62.1	0.009
LA [deg]	420	13.0	3.43	12.3	5.78	22.1	6.90	<.001
ROM [deg]	420	70.6	6.87	70.3	8.84	93.8	50.3	0.708
RV [deg/s]	420	69.8	7.00	69.6	8.52	93.7	48.3	0.209
PPT [s]	361	2.53	0.814	2.40	0.82	7.80	1.52	<.001
OR [deg]	420	33.4	7.45	33.0	9.86	60.3	12.2	0.009

Table 4.11: Descriptive statistic of OPT values for frontal execution. P-values less than 0.05 indicate a non-normal distribution. N indicates the number of parameters considered.

Frontal MP

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
HA [deg]	420	96.1	10.1	95.4	13.8	131	71.3	0.06
LA [deg]	420	21.9	4.04	21.2	6.13	34.9	12.8	<.001
ROM [deg]	420	74.2	10.7	73.3	14.8	110	49.0	0.018
RV [deg/s]	420	69.1	14.2	68.4	18.1	108	13.3	0.008
PPT [s]	361	2.53	0.820	2.38	0.80	7.74	1.42	<.001
OR [deg]	420	46.1	10.1	45.4	13.8	81.4	21.3	0.06

Table 4.12: Descriptive statistic of MP values for frontal execution. P-values less than 0.05 indicate a non-normal distribution. N indicates the number of parameters considered.

Table 4.9 and Table 4.10 provide the descriptive statistics for parameters extracted from the lateral tasks, while Tables 4.11 and 4.12 pertain to frontal tasks.

A greater sample availability is observed in the lateral raises compared to the frontal ones, as some repetitions in the frontal task were excluded due to excessive noise.

In general, the descriptive statistics are consistent with the findings in the aggregate parameters analysis, confirming that the values extracted with the OPT system exhibit lower means than those computed with the MP system. This supports the tendency of the MP model to overestimate the angular measures. The same trend is noted in the variability measures, expressed by the Interquartile range (IQR) and standard deviation (SD), with an exception in the HA parameter in the lateral raises, where variability is greater in the OPT signal (SD: 7.44° vs 7.25°, IQR: 11.2° vs 10.5°). This discrepancy may be due to the heterogeneity of the subjects' movement patterns, that the OPT system is able to detect due to its higher sensitivity.

The temporal parameters reveal highly comparable values between the two systems, under both operating conditions, as also highlighted by the boxplots in Figure 4.9. The same outliers are observed in both datasets, corresponding to the alternating raises, that involve longer temporal intervals compared to other movements.

Most of the parameters present a non-normal distribution, as indicated by the p-value of the Shapiro-Wilk test (SW), falling below the significance threshold

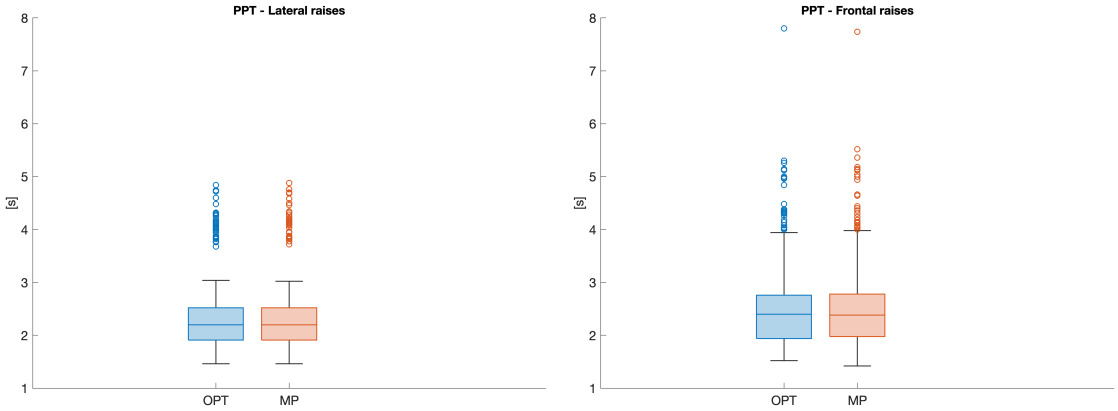

Figure 4.9: Boxplots of PPT distribution for lateral (left) and frontal (right) raises. In blue the boxplot related to the OPT signal, in red this related to MP signal

(0.05). Exceptions include the ROM and RV parameters in the frontal raises of the OPT signal and, marginally, the HA and RV in the MP signal, where the p-value is slightly higher than the threshold (p=0.06).

Overall, the mean and median values are closely aligned in both lateral and frontal raises, suggesting symmetrical and balanced distributions.

As previously observed in the analysis of aggregate parameters, the variability and dispersion indices are more marked in the parameters computed from MP compared to those extracted from the OPT signals. This effect is more evident in frontal tasks, where the variability is higher.

The RV parameter demonstrates increased variability in the MP system. Although this variability is relatively high, it can be considered acceptable, given that RV is a parameter strongly influenced by both movement heterogeneity and by the accuracy of the segmentation algorithm.

In conclusion, the findings of the preliminary analysis conducted on the mean values are confirmed:

- The temporal parameters show excellent reliability.
- The angular and velocity parameters exhibit greater variability, in particular in the frontal raises.

Regarding the correlation analysis, the Table 4.13 reports the correlation coefficients along with their relative p-values for each extracted parameter, for the lateral and frontal raises.

In this case, correlation analysis was performed using Spearman's rho coefficient, as at least one of the two distributions did not meet the assumption of normality.

Correlation analysis

Parameter	Spearman's rho (Lateral)	P-value	Spearman's rho (Frontal)	P-value
HA	0.693***	<.001	0.521***	<.001
LA	−0.360***	<.001	−0.275***	<.001
ROM	0.749***	<.001	0.313***	<.001
RV	0.780***	<.001	0.346***	<.001
PPT	0.993***	<.001	0.971***	<.001
OR	0.693***	<.001	0.521***	<.001

Table 4.13: Spearman correlation coefficients for all parameters in both lateral and frontal raises. *Note.* * $p < .05$, ** $p < .01$, *** $p < .001$.

In general, stronger and more statistically significant correlations were observed during lateral raises compared to frontal ones.

The angular parameter LA presents a slight negative correlation, respectively equal to -0.360 in lateral raises and -0.275 in frontal raises. The presence of a negative correlation indicated that as the values measured by OPT decrease, those measured by MP increase. This highlights a limited stability in tracking the minimum points, which negatively affects the segmentation algorithm's accuracy in detecting minima correctly. For this reason, the positions of the minima were excluded for the calculation of the temporal parameters.

The HA parameter shows a moderate correlation in both lateral (0.693) and frontal raises (0.521).

The ROM parameter reveals a marked difference in correlation coefficients between lateral and frontal raises: it shows a strong correlation in the lateral task (0.749), and a weaker in the frontal task (0.313). A similar trend is evident in the velocity parameter RV, where the correlation coefficient is 0.780 in lateral raises and 0.346 in frontal raises.

These discrepancies may be caused by the instability of the tracking at the minimum points in the frontal trajectories, or by a reduced accuracy of the segmentation algorithm under less regular signal morphologies.

However, the temporal parameters exhibit a very high correlation values, close to one in both tasks: 0.993 for lateral raises and 0.971 for frontal raises. These outcomes further confirm the robustness and reliability of temporal parameters, independently from the direction of the movement.

Furthermore, it is observed that OR presents the same values as HA, as these two measures differ only by a constant.

In conclusion, in lateral raises a good correlation is evident in all the parameters except for LA. By contrast, during frontal raises, most parameters show only moderate correlations, with the exception for the temporal parameter.

The high correlation of PPT parameter, calculated with the time position of the maximum, suggests a greater synchronization and stability of the signal in the tracking of movements with wide angular excursions. This is also suggested by the slight correlation of LA parameter.

For a more in-depth evaluation of the agreement between the two systems, the Bland-Altman analysis was further conducted.

Figure 4.10 shows the Bland-Altman graphs for the LA parameter, corresponding to the lateral (left) and frontal (right) execution.

Almost the entire observations are contained within the Limits of Agreement (LoA), which exhibit significant variability in both conditions (Lateral: [-18.16°, 3.29°], Frontal: [-20.71°,2.79°]). These wide intervals indicate a significant dispersion of this parameter, regardless of the execution type, confirming the low correlation found in the previous analysis as well as the pronounced differences in mean values observed both in the aggregate (i.e., the averaged parameters analysis) and in the single repetitions analysis.

The mean bias is negative in both conditions (-7.43° in lateral raises, -8.96° in frontal raises), demonstrating that MP tends to overestimate this parameter compared to OPT, as observed previously.

The negative slope of the regression line reveals a proportional bias: as the mean value of the parameter increases, the difference between the two measures decreases.

The analysis of the individual tasks reveals a consistent distribution pattern. In lateral raises only few values fall outside the LoA, in particular, those related to simultaneous and alternating raises. In frontal raises, the few outliers are distributed across all task types, probably due to the more irregular morphology in the associated trajectories.

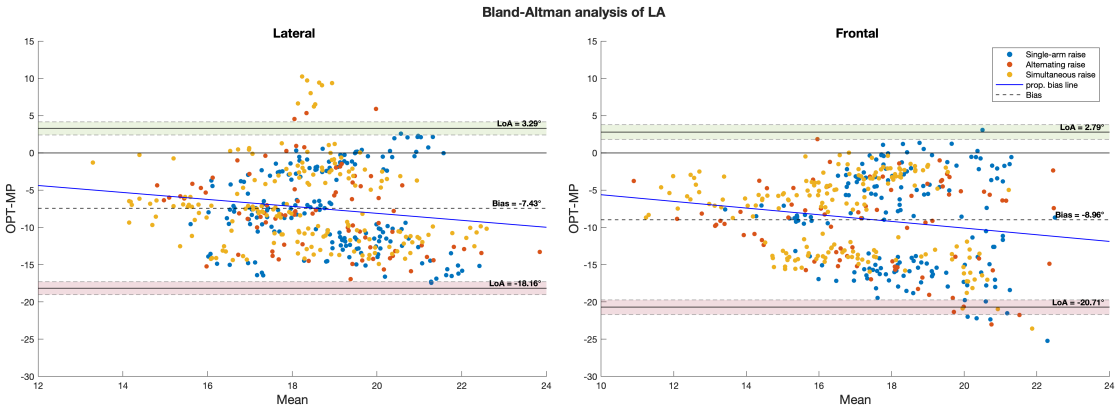

Figure 4.10: Bland-Altman analysis of LA parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

Figure 4.11 presents the Bland-Altman analysis conducted on the ROM parameter.

The bias in the lateral tasks is 2.03° while in the frontal tasks it is -3.56°,

revealing that the MP model tends to systematically underestimate ROM in the lateral raises and overestimate it in the frontal raises. These findings are in line with what was already observed in the aggregate parameter analysis. The agreement interval is narrower in the lateral raises [-9.16°, 13.22°] compared to the frontal ones [-23.53°, 16.42°], revealing a higher coherence between systems in the lateral executions.

In both cases, the regression line shows a negative slope, indicating that as the mean value of the two measures increases, the difference between them tends to decrease. This effect is clear in the single-arm task in lateral raises and for simultaneous task in frontal executions.

From the analysis of the individual tasks, it is evident that the outliers in the lateral raises are mainly associated with the single-arm lift, suggesting lower agreement between systems in this specific exercise. On the other hand, simultaneous movements exhibit a more clustered distribution around the bias, indicating a greater accuracy in the ROM computed from this task. This may be due to a enhanced tracking stability favoured by the simultaneous limb movements.

In the frontal raises, a uniform distribution is evident across different tasks: a greater concentration of the simultaneous measures falls below zero. This suggests that, especially for this task, the MP model tends to overestimate compared to the reference system.

The alternate raises appear more dispersed in both executions: this may be due to the greater variability of the trajectories, influenced by the transient wave problem.

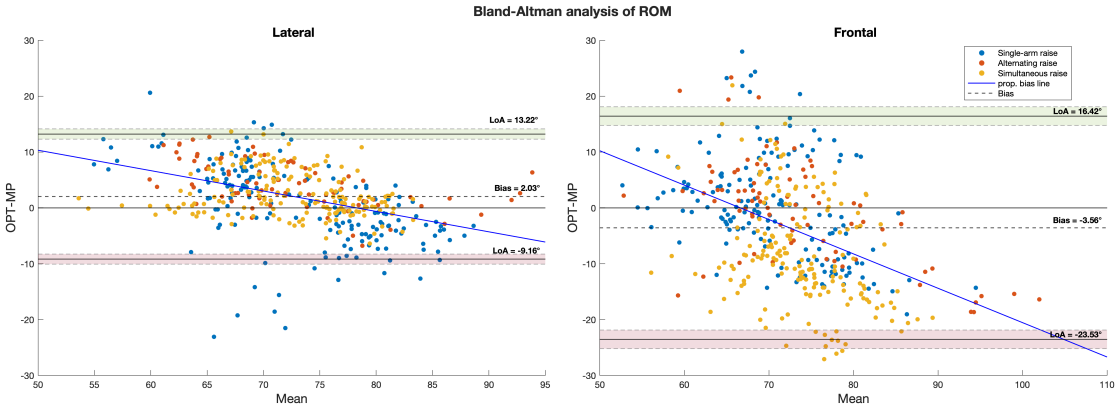

Figure 4.11: Bland-Altman analysis of ROM parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

The Bland-Altman analysis of the PPT parameter is reported in Figure 4.12 and demonstrated a high level of agreement between the two systems.

In lateral raises, the limits of agreement are extremely narrow [-0.09 s, 0.09 s] and no bias is observed, indicating a near-perfect temporal alignment between the two systems. In frontal raises, the LoA are slightly wider, defining an interval between -0.23 s and 0.24 s, which is completely acceptable in the application context. The bias is very low, equal to 0.01 s.

It is evident that the alternating raises exhibit higher average values, which is consistent with the nature of the movement (the temporal distance between adjacent peaks is greater in alternating movements).

Although a few outliers are observed, they are not problematic, as they present differences of just over one-tenth of a second, which are compatible with the tolerance margins expected for the practical use of the exergame.

In conclusion, the Bland-Altman analysis confirms the robustness and reliability of the PPT parameter, showing a high agreement between the systems. This result is further justified by the strong correlation observed.

Conversely, the angular parameter LA presents, in both executions, a poor agreement, characterized by a non-negligible bias and a very wide agreement interval, indicating a greater variability of the tracking in minimum points.

Finally, the ROM demonstrates a good agreement in lateral raises, as also confirmed by the correlation coefficient, whereas only moderate agreement in frontal raises is found, where the dispersion is greater.

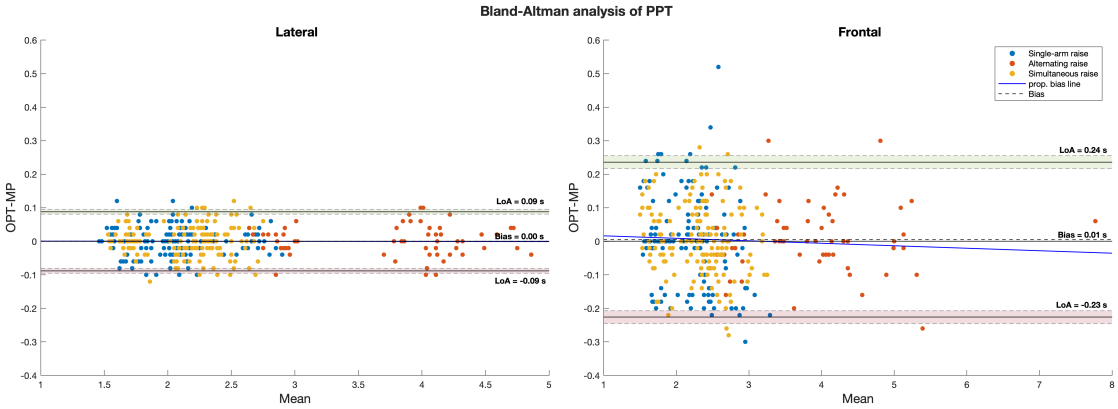

Figure 4.12: Bland-Altman analysis of PPT parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

Bilateral parameter analysis

The tables below report the descriptive statistics related to the bilateral parameters, which include the mean, standard deviation, median, interquartile range (IQR), maximum, minimum values and the significance of the normality test.

Lateral OPT

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
ΔHA [deg]	130	5.67	4.17	4.97	5.72	15.3	0.001	<.001
ΔLA [deg]	130	2.28	1.48	2.05	2.09	6.76	0.02	<.001
ΔROM [deg]	130	4.80	3.56	3.98	5.08	13.50	0.06	<.001
IPID [s]	110	0.0495	0.0572	0.02	0.04	0.320	0.00	<.001
PTL [s]	130	0.597	0.890	0.02	1.42	2.48	0.00	<.001
ΔRV [deg/s]	130	5.05	4.21	4.22	5.49	22.8	0.01	<.001

Table 4.14: Descriptive statistic of OPT values for lateral execution. P-values less than 0.05 indicate a non-normal distribution.

Lateral MP

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
ΔHA [deg]	130	4.59	4.16	3.37	6.48	19.5	0.001	<.001
ΔLA [deg]	130	2.62	1.49	2.75	2.23	5.92	0.00	0.001
ΔROM [deg]	130	4.67	3.53	3.85	4.32	16.90	0.11	<.001
IPID [s]	110	0.0735	0.065	0.06	0.08	0.3	0.00	<.001
PTL [s]	130	0.608	0.878	0.06	1.42	2.48	0.00	<.001
ΔRV [deg/s]	130	5.05	4.12	4.56	6.21	22.4	0.03	<.001

Table 4.15: Descriptive statistic of MP values for lateral execution. P-values less than 0.05 indicate a non-normal distribution.

Frontal OPT

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
ΔHA [deg]	129	2.56	2.81	1.76	2.69	21.0	0.01	<.001
ΔLA [deg]	129	2.02	1.18	1.87	1.48	5.29	0.03	0.003
ΔROM [deg]	129	2.96	2.53	2.40	2.75	17.9	0.00	<.001
IPID [s]	109	0.0462	0.0740	0.02	0.04	0.480	0.00	<.001
PTL [s]	129	0.606	0.922	0.02	1.44	2.70	0.00	<.001
ΔRV [deg/s]	129	3.36	2.68	2.78	3.01	18.2	0.00	<.001

Table 4.16: Descriptive statistic of OPT values for frontal execution. P-values less than 0.05 indicate a non-normal distribution.

Frontal MP

Parameter	N	Mean	SD	Median	IQR	Max	Min	SW p-value
ΔHA [deg]	129	8.03	5.30	6.85	5.92	25.7	0.130	<.001
ΔLA [deg]	129	2.30	1.80	2.15	2.73	8.65	0.01	<.001
ΔROM [deg]	129	7.75	5.47	7.20	6.22	23.1	0.05	<.001
IPID [s]	109	0.133	0.110	0.120	0.14	0.580	0.00	<.001
PTL [s]	129	0.636	0.896	0.10	1.36	2.90	0.00	<.001
ΔRV [deg/s]	129	9.20	7.45	7.15	7.85	38.6	0.170	<.001

Table 4.17: Descriptive statistic of MP values for frontal execution. P-values less than 0.05 indicate a non-normal distribution.

Tables 4.14 and 4.15 present the descriptive statistics of the parameters computed respectively for the MP and OPT systems during lateral raises, while Tables 4.16 and 4.17 report the same analysis for the frontal raises.

In general, in lateral raises, the mean and median values are closely between the two systems. The most marked difference is observed in the ΔHA parameter, whose mean value is 5.67° for the OPT and 4.59° for MP. Both IQR and SD are relatively high compared to the other parameters, suggesting a discrete variability. The greater mean and median values observed in OPT may be attributable to a higher sensitivity of the optoelectronic system, in detecting intra-task variability. This trend is also observed in the frontal raises, where it is even more pronounced.

The ΔLA parameter exhibits comparable mean, median, and variability values between the two systems in both lateral and frontal raises. This evidence also emerged in the aggregate parameters analysis, suggesting that the exergame's tracking detects symmetric minima between both arms of the same subject. This may indicate that the tracking is influenced by the anthropometric characteristics of the subject.

The ΔROM parameter presents highly comparable values between MP and OPT in lateral raises, in terms of both mean and variability. In frontal raises, however, a more marked difference in the mean value (MP: 7.75° OPT: 2.96°) is observed.

The temporal parameters remain consistent across the systems, confirming what was observed from the aggregate parameters analysis. The PTL parameter reveals a marked variability in both systems and in both execution modes. This is expectable, as alternating raises are physiologically characterized by longer inter-peak intervals than the other tasks.

The ΔRV parameter yields almost equal mean values in lateral raises, while in frontal raises the difference is more marked (OPT: 3.36 °/s, MP: 9.20 °/s). This can be explained by the more irregular morphology that characterized the frontal angular trajectories.

In general, none of the analyzed parameters follow a normal distribution, as evidenced by the p-values of the Shapiro-Wilk test; therefore, the correlation between the parameters was evaluated using the Spearman coefficient.

Table 4.18 reports the Spearman correlation coefficients for lateral and frontal raises along with their relative p-values for each extracted parameter.

Correlation analysis

Parameter	Spearman's rho (Lateral)	P-value	Spearman's rho (frontal)	P-value
ΔHA	0.473***	<0.001	0.316***	<.001
ΔLA	0.251**	0.004	0.093	0.293
ΔROM	0.459***	<0.001	0.119	0.181
IPID	0.459***	<0.001	0.322***	<.001
PTL	0.768***	<0.001	0.670***	<.001
ΔRV	0.270**	0.002	0.151	0.088

Table 4.18: Spearman correlation coefficients for all parameters in both lateral and frontal raises. *Note.* * $p < .05$, ** $p < .01$, *** $p < .001$.

In general, the correlation coefficients prove to be higher in lateral raises compared to frontal ones. However, excluding the temporal parameter PTL, none of the parameters exhibit strong correlation.

The ΔHA parameter reveals a moderate correlation in lateral task (0.473) and a limited correlation in frontal task (0.316). On the other hand, ΔLA demonstrates a negligible correlation both in lateral (0.251) and frontal raises (0.093). In the latter case, it fails to reach significance, as evidenced by a p-value of 0.293.

The ΔROM exhibits a moderate correlation in lateral executions (0.459) and poor in frontal ones (0.119, p: 0.181).

Temporal parameters generally show stronger correlations than angular parameters, with similar values across frontal and lateral executions. IPID shows a moderate correlation in the lateral task (0.459) and a weaker correlation in the frontal ones (0.322). PTL emerges for being the only parameter to have a strong correlation in both executions (0.768 in the lateral raises and 0.670 in the frontal raises).

The ΔRV parameter exhibits a negligible correlation both in the lateral (0.270) and in the frontal raises (0.151, p: 0.088).

In summary, these results suggest that the temporal parameters are characterized by a higher correlation than the others. Moreover, the agreement between systems is higher in the lateral raises than in the frontal raises.

The lower correlation values observed for these bilateral parameters, compared to the unilateral ones, may be attributed to the sum of several factors: the variability introduced by the system's tracking, the segmentation algorithm and the presence of an intra-subject variability (not detectable by the exergame).

The Bland-Altman analysis conducted on the ΔROM parameter (Figure 4.13) presents notable differences between the measurements related to lateral and frontal raises. In lateral raises, the bias is 0.13°, while in frontal raises it assumes a value of -4.79°, revealing a slight overestimation of the parameter. The LoA for lateral raises defined the interval [-7.12°, 7.38°], while those related to frontal raises ranged from -14.97° to 5.38°.

In frontal raises, the presence of a significant proportional bias is evident: as the mean value of the two measurements increases, the difference between them progressively decreases.

In lateral raises, the dispersion of the measurements is similar between simultaneous and alternating executions. On the other hand, during frontal raises, simultaneous measurements tend to be clustered around the mean value 5°, unlike those relating to alternating raises that present greater deviation. This phenomenon may be explained by the presence of the convex wave transient, which could influence the accuracy of the segmentation algorithm.

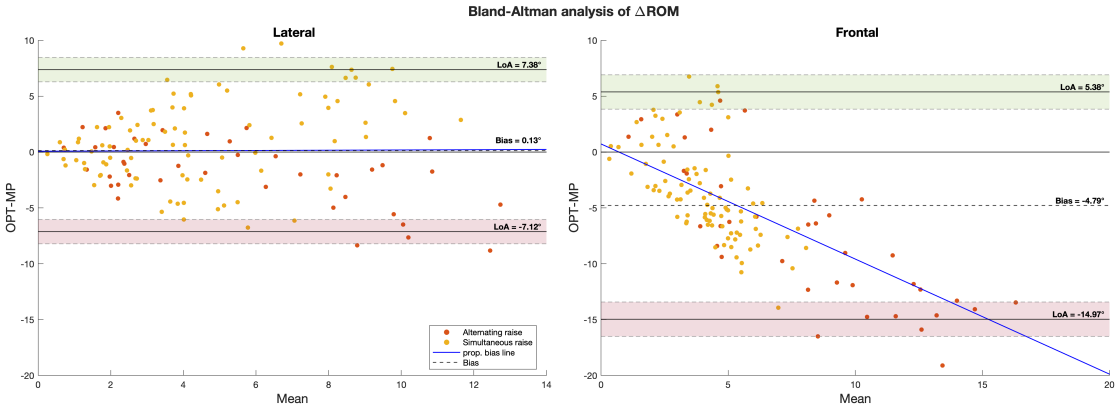

Figure 4.13: Bland-Altman analysis of ΔROM parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: red for alternating raises, and yellow for simultaneous ones.

Figure 4.14 displays the Bland-Altman plots for the ΔRV parameter which reveal a pattern similar to that observed for ΔROM parameter. In Lateral raises, the bias is insignificant and the LoA are symmetric [-10.22°/s, 10.22°/s], while in frontal raises asymmetric Limit of Agreement [-20.24°/s, 8.55°/s] and a negative systemic bias (-5.84°/s) are observed. This indicates a MP tendency to overestimate the RV values regarding OPT.

Even in this case, a negative slope can be seen in the regression line, indicating the presence of a proportional bias as observed in ΔROM parameter. This may be attributed to the segmentation algorithm, as the RV parameter is derived from ROM.

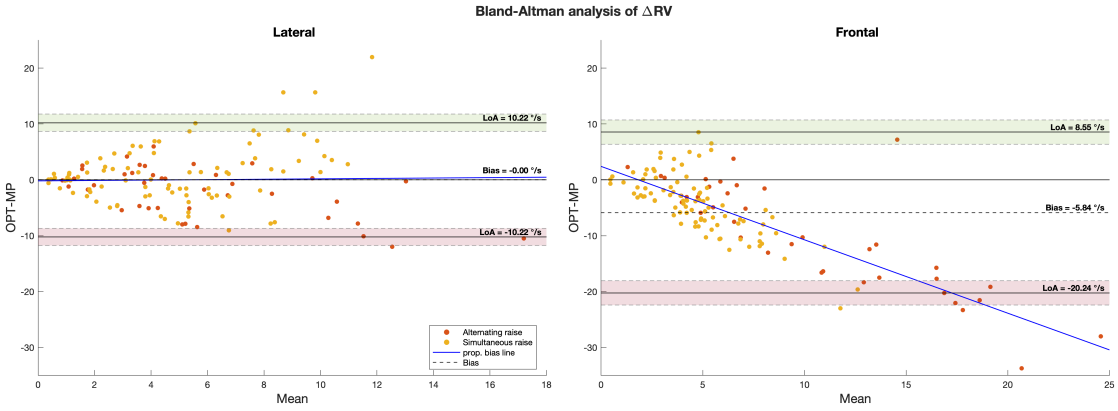

Figure 4.14: Bland-Altman analysis of ΔRV parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: red for alternating raises, and yellow for simultaneous ones.

Figure 4.15 displays the Bland-Altman plots for the PTL parameter.

In lateral raises, a systematic bias of -0.01 s occurs, with limits of agreement between -0.09 s and 0.07 s. In frontal raises, higher variability is noted: the bias increases to -0.03 s, and the LoA range from -0.23 s to 0.17 s.

Analysing the individual tasks, it is evident that the values associated with simultaneous raises are located near zero, while those corresponding to alternating raises are more dispersed. This behaviour, presents in both lateral and frontal raises, confirms the capability of the parameter to effectively distinguish between different types of movement, each characterized by different times.

Despite the presence of some outliers, the agreement between the systems remains good. It should be emphasized that these parameters become clinically significant when they present values in the order of a second.

In conclusion, even in bilateral parameters, a higher concordance between systems is evident in lateral raises, compared to frontal ones, as suggested by both the correlation and Blan-Altman analysis.

Overall, except for the temporal parameter, the outcomes assessed did not show a strong agreement between the two systems. This suggest that the MP system may is not sensitive enough to accurately detect parameters derived from values close to zero, especially in angular parameters, where the variability introduced by the segmentation algorithm is more pronounced.

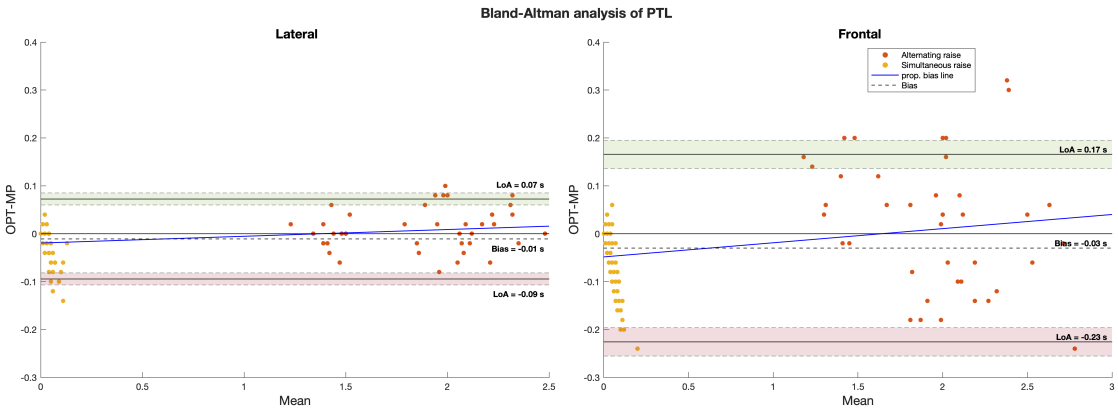

Figure 4.15: Bland-Altman analysis of PTL parameter, conducted for lateral (left) and frontal (right) raises. Each task is color-coded: red for alternating raises, and yellow for simultaneous ones.

4.3.2 OPT-based segmentation

This section reports the most significant results from the analysis of the parameters computed through the segmentation based on the OPT signal (maxima and minima position obtained from the reference signal), to quantify the impact of the segmentation strategy on the validation results. Only angular parameters are presented, as the temporal ones are identical to those previously obtained from the OPT analysis. For simplicity, parameters extracted from the MP signal using the standard segmentation algorithm are defined as "*MP1*", while those computed by applying the OPT-based segmentation are referred to as "*MP2*".

Unilateral parameters analysis

An analysis of the distributions (Figure 4.16 and 4.17) shows no great differences between the parameters computed using MP1 compared to those calculated with MP2, both in lateral and frontal raises. In general, the interquantile range tends to be slightly lower for MP2, while the overall dispersion appears greater, as indicated by the extended whiskers in the boxplots.

This behaviour is plausible, as the error introduced by the segmentation on the angular signals obtained from the exergame is limited to few degrees; therefore, it does not influence parameters whose values are on the order of tens of degrees.

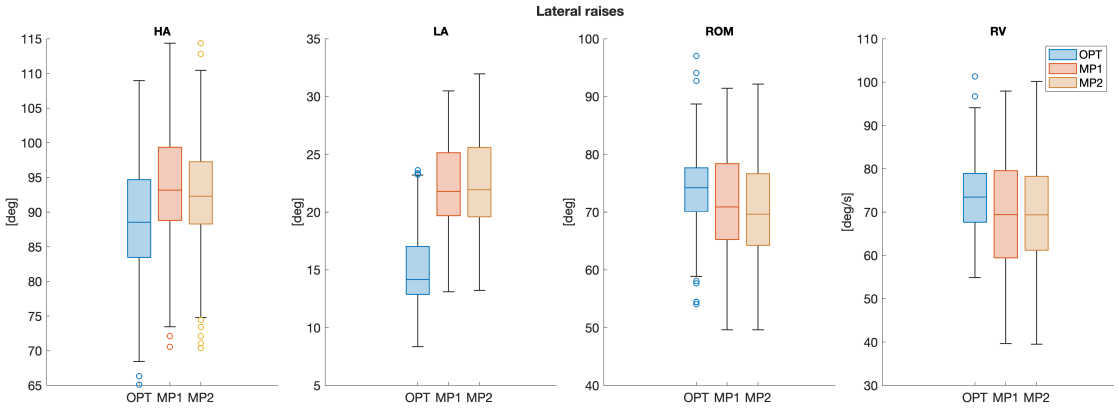

Figure 4.16: Boxplot of angular parameters for lateral raises. The distribution of parameters extracted from the OPT signal are displayed in blue, those from MP1 in red, and those from MP2 in yellow.

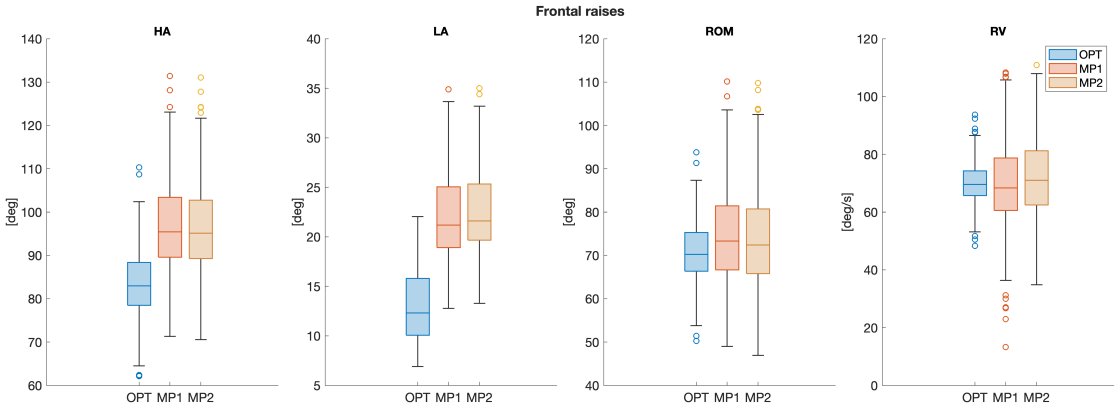

Figure 4.17: Boxplot of angular parameters for frontal raises. The distribution of parameters extracted from the OPT signal are displayed in blue, those from MP1 in red, and those from MP2 in yellow.

The analysis of Spearman correlation coefficients (Tables 4.19 and 4.20) reports that, generally, the correlation is higher in the parameters calculated using MP2, in both lateral and frontal raises, with two exceptions:

- In lateral raises, the negative correlation slightly decreases for the LA parameter (MP1: -0.360, MP2: -0.314).
- In frontal raises, the correlation for the HA moderately decreases (MP1: 0.521, MP2: 0.519).

Correlation analysis - Lateral raises

Parameter	Spearman's rho MP1	p-value	Spearman's rho MP2	p-value
HA	0.693***	<.001	0.769***	<.001
LA	−0.360***	<.001	−0.314***	<.001
ROM	0.749***	<.001	0.849***	<.001
RV	0.780***	<.001	0.893***	<.001

Table 4.19: Spearman correlation coefficient for all unilateral parameters in both MP1 and MP2 during lateral raises. *Note.* * $p < .05$, ** $p < .01$, *** $p < .001$.

Correlation analysis - Frontal raises

Parameter	Spearman's rho MP1	p-value	Spearman's rho MP2	p-value
HA	0.521***	<.001	0.519***	<.001
LA	−0.275***	<.001	−0.284***	<.001
ROM	0.313***	<.001	0.366***	<.001
RV	0.346***	<.001	0.379***	<.001

Table 4.20: Spearman correlation coefficient for all unilateral parameters in both MP1 and MP2 during frontal raises. * $p < .05$, ** $p < .01$, *** $p < .001$.

The overall increase in correlation coefficients suggests that a relevant part of the variability between the two systems may be attributed to the segmentation algorithm. The most pronounced improvements are observed in lateral raises, probably because the angular trajectories have a similar morphology to those computed with the optoelectronic system. Under these conditions, it is more probable that the values extracted at the same time points in the realigned signals corresponds to equal ranks in the correlation analysis.

In contrast, in frontal raises, the improvements are less evident, probably due to the different morphology of the frontal trajectories.

From the Bland-Altman graphs it is possible to deepen the analysis and observe how much the segmentation influences the agreement between the two systems.

The ROM parameter, which exhibits a moderate increase in the correlation coefficient, shows an improvement in the agreement, especially in the lateral raises (Figure 4.18). In contrast, in the frontal raises (Figure 4.19) the improvements appear negligible. Notably, in lateral raises, the limits of agreement are reduced: for MP1 the LoA define a range between -9.16° and 13.22°, while on MP2 this range narrows to [-5.03°, 11.99°], indicating a reduced dispersion in the measurements.

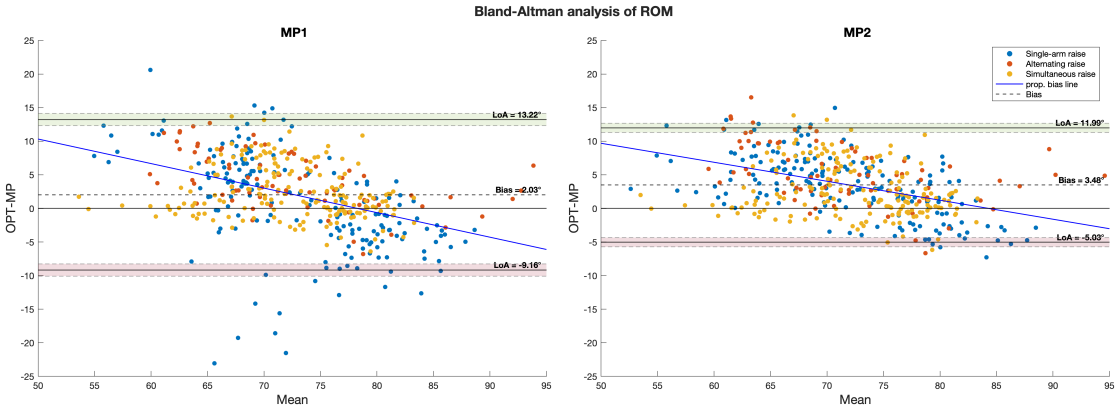

Figure 4.18: Bland-Altman analysis of ROM parameter, conducted for MP1 and MP2 during **lateral raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

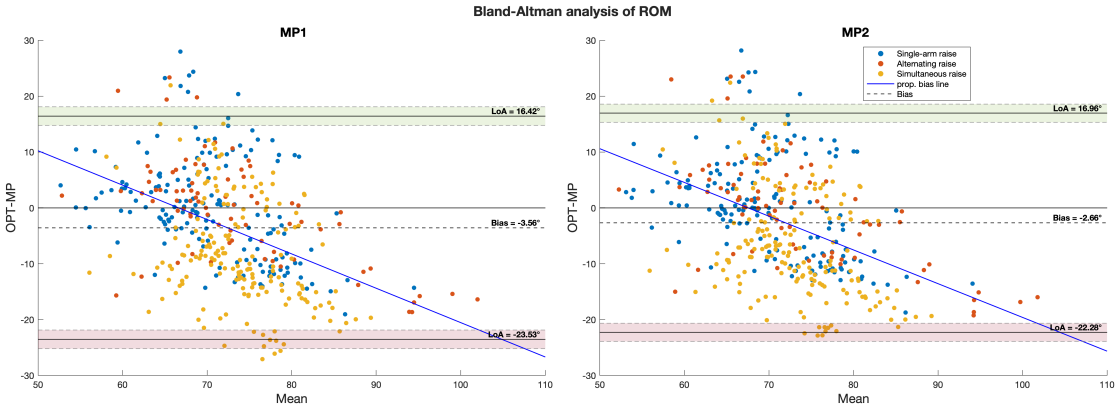

Figure 4.19: Bland-Altman analysis of ROM parameter, conducted for MP1 and MP2 during **frontal raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

A comparable behaviour is observed in the RV parameter. In the lateral raises (Figure 4.20), the agreement between the two systems increases more evidently than in the frontal raises (Figure 4.21), which exhibit only a limited improvement.

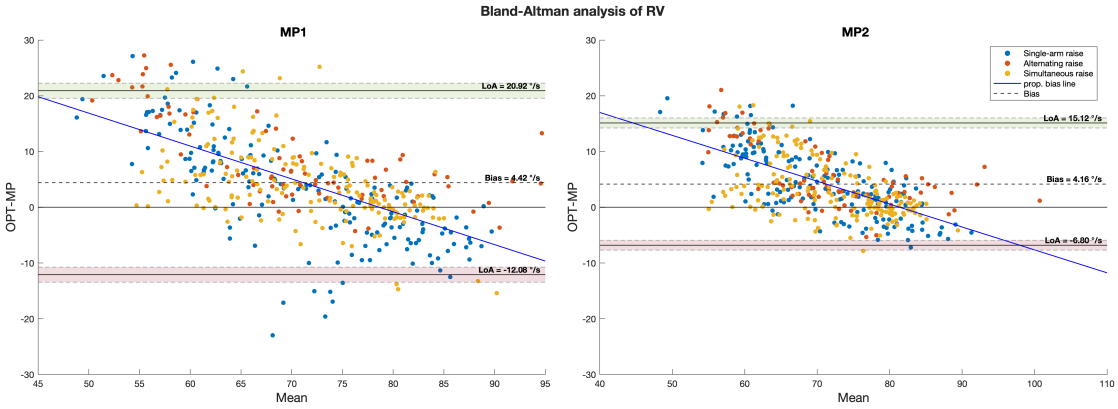

Figure 4.20: Bland-Altman analysis of RV parameter, conducted for MP1 and MP2 during **lateral raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

Analyzing the LA parameter (Figure 4.22 and Figure 4.23), which presents a weak correlation in both MP1 and MP2, it is noted that the agreement between the two systems does not enhance significantly. This suggests that, in regions where the tracking is unstable (i.e., at the minimum or transition points), the variability introduced by the segmentation is negligible compared to the intrinsic variability of the signal. This phenomenon is particularly evident in the frontal raises, in which the morphology of the signal is more irregular compared to that observed in lateral raises.

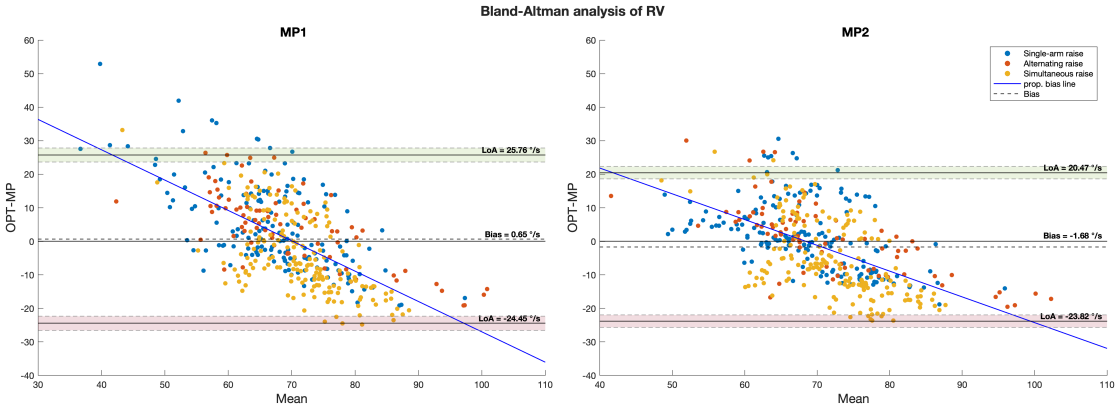

Figure 4.21: Bland-Altman analysis of RV parameter, conducted for MP1 and MP2 during **frontal raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

In conclusion, the standard segmentation introduces variability in the measurements. This variability is more significant in the lateral raises than in the frontal raises, in which the intrinsic variability of the signal overwhelms that introduced by the segmentation.

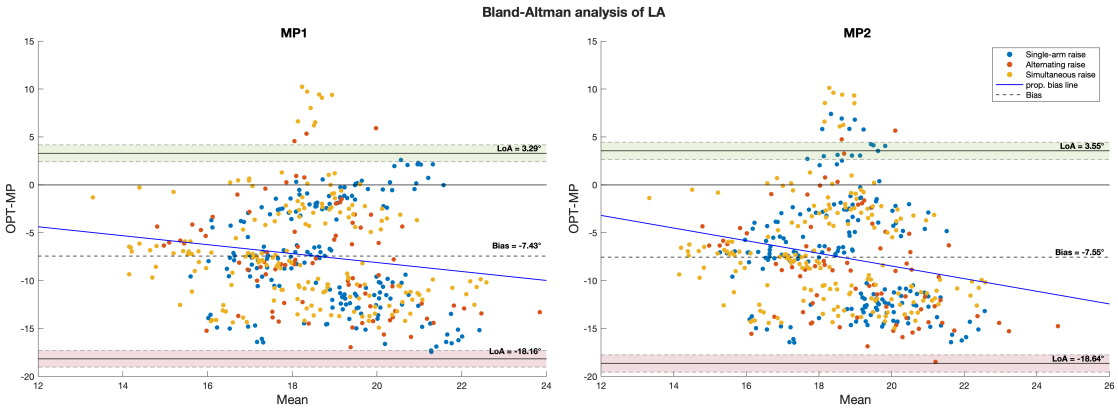

Figure 4.22: Bland-Altman analysis of LA parameter, conducted for MP1 and MP2 during **lateral raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

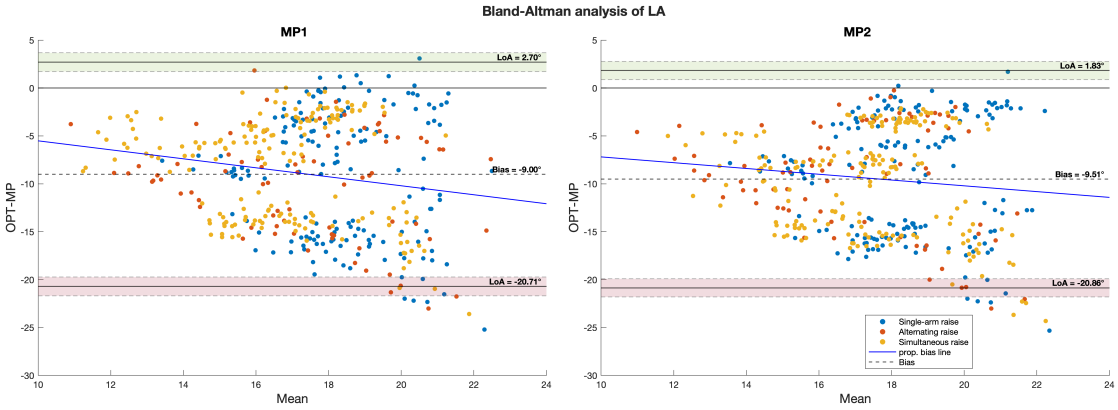

Figure 4.23: Bland-Altman analysis of LA parameter, conducted for MP1 and MP2 during **frontal raises**. Each task is color-coded: blue for single-arm raise, red for alternating raises, and yellow for simultaneous ones.

Bilateral parameters analysis

This section presents the bilateral parameters analysis, discussing on the effect of the segmentation algorithm.

Figures 4.24 and 4.25 present boxplots illustrating the distributions of bilateral parameters computed on both MP1 and MP2. In lateral raises, the parameters

exhibit a higher variability than in frontal raises. This is also observed in OPT data, suggesting the presence of intra-subject variability during bilateral tasks execution.

The distributions of MP1 and MP2 are generally comparable across all parameters, for both executions, with similar variability and comparable median values. The only exception is represented by ΔLA, which reveals a reduced interquartile variability but a higher median value than both MP1 and OPT.

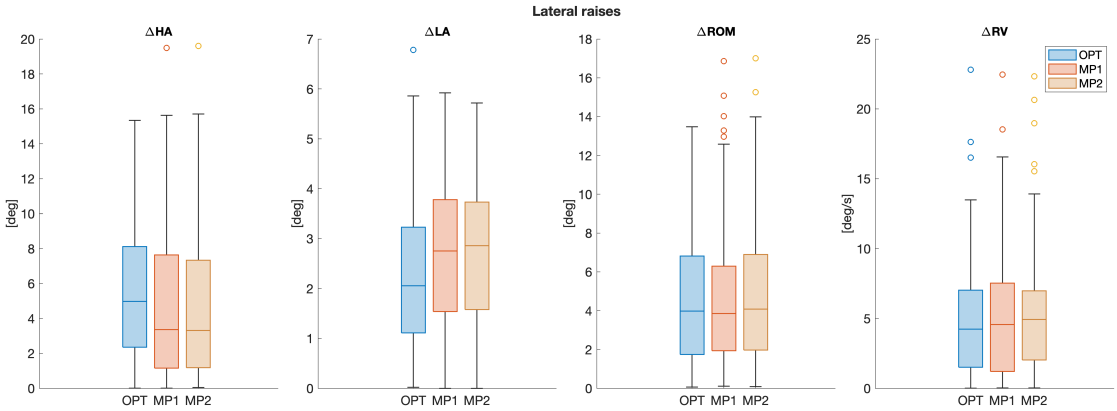

Figure 4.24: Boxplot of bilateral angular parameters for lateral raises. The distribution of parameters extracted from the OPT signal are displayed in blue, those from MP1 in red, and those from MP2 in yellow.

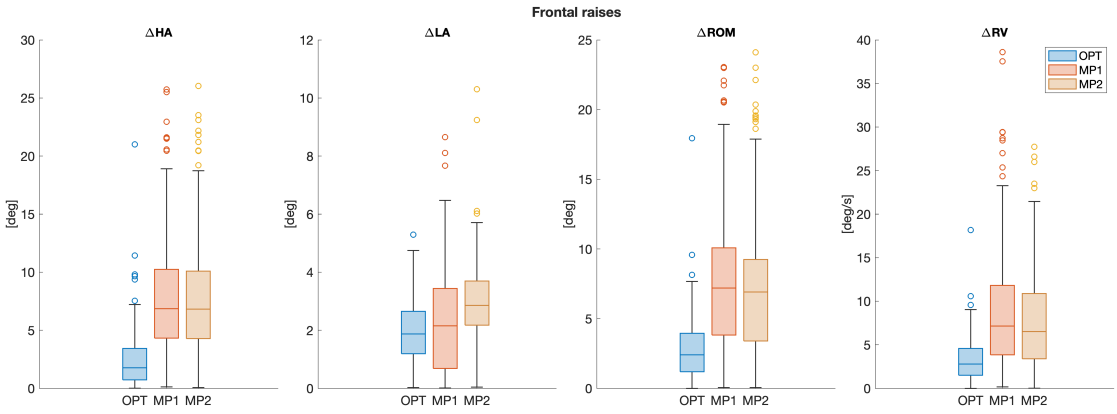

Figure 4.25: Boxplot of bilateral angular parameters for frontal raises. The distribution of parameters extracted from the OPT signal are displayed in blue, those from MP1 in red, and those from MP2 in yellow.

Tables 4.21 and 4.22 provide the Spearman correlation coefficients of the parameters calculated from MP1 and MP2, respectively related to lateral and frontal raises.

Correlation analysis -Lateral raises

Parameter	Spearman's rho MP1	p-value	Spearman's rho MP2	p-value
ΔHA	0.473***	<0.001	0.470***	<.001
ΔLA	0.251**	0.004	0.301***	<.001
ΔROM	0.459***	<0.001	0.470***	<.001
ΔRV	0.270**	0.002	0.449***	<.001

Table 4.21: Spearman correlation coefficient for all bilateral parameters in both MP1 and MP2 during lateral raises. * $p < .05$, ** $p < .01$, *** $p < .001$.

Correlation analysis -Frontal raises

Parameter	Spearman's rho MP1	p-value	Spearman's rho MP2	p-value
ΔHA	0.316***	<.001	0.324***	<.001
ΔLA	0.093	0.293	0.301***	<.001
ΔROM	0.119	0.181	0.012	0.888
ΔRV	0.151	0.088	0.115	0.195

Table 4.22: Spearman correlation coefficient for all bilateral parameters in both MP1 and MP2 during frontal raises. * $p < .05$, ** $p < .01$, *** $p < .001$.

No generalized trend is evident, as some parameters exhibit an increase in the correlation coefficient when computed with MP2, while other show a reduction, in both lateral and frontal raises.

A moderate enhancement is observed in the ΔRV parameter during lateral raises, whose coefficient increases from 0.270 (MP1) to 0.449 (MP2).

Another improvement can be noted in the ΔLA parameter during frontal raises. In this case, the correlation coefficient is 0.093 in MP1 and increases to 0.301 in MP2.

In the frontal raises, the ΔROM parameter presents limited correlation values for both MP1 and MP2, probably because the distribution of the OPT signal is clustered, while that of MP1 and MP2 is dispersed.

In general, the negligible correlation values observed across both segmentation methods may be due to the bilateral parameter nature, which are computed as the difference between the values of two limbs. Since the analysis was performed on healthy subjects, where the inter-limbs symmetry is generally high, the differences tend to be negligible, near zero. Therefore, such parameters are sensitive to errors induced by tracking. In similar conditions, the informative component of the signal, reflecting inter-limb differences, is deleted, and only the intrinsic variability remains. This explains the lower correlation values, especially in the parameters that have lower values. Consequently, the variability introduced by the standard segmentation is negligible in this case.

The Bland-Altman analysis confirms the observations emerged from the correlation analysis.

For the ΔROM parameter (Figure 4.26 and 4.27), the agreement between the two systems is not significantly influenced by the variability introduced by the

segmentation algorithm, neither in the frontal nor in the lateral raises. The LoA remain almost unchanged for both MP1 and MP2, delineating approximately the same range of variability. Similarly, the bias is nearly the same in both movement conditions.

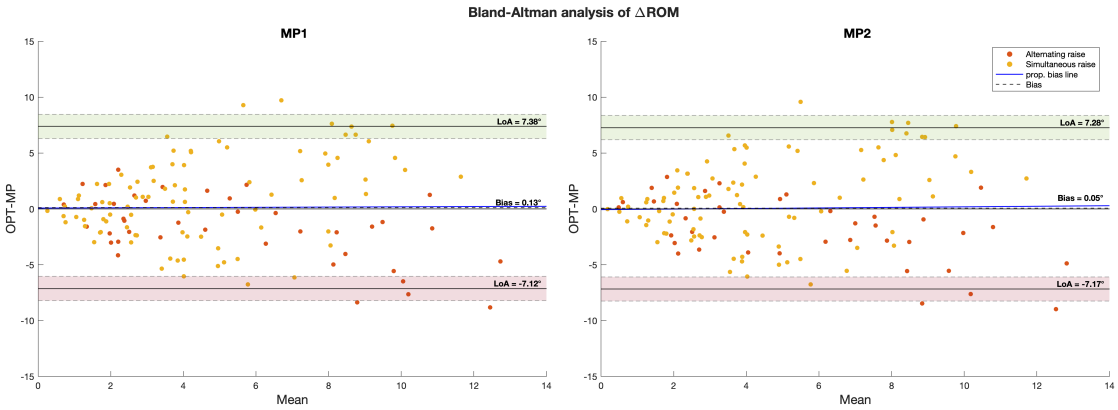

Figure 4.26: Bland-Altman analysis of ΔROM parameter, conducted for MP1 and MP2 during **lateral raises**. Each task is color-coded: red for alternating raises and yellow for simultaneous ones.

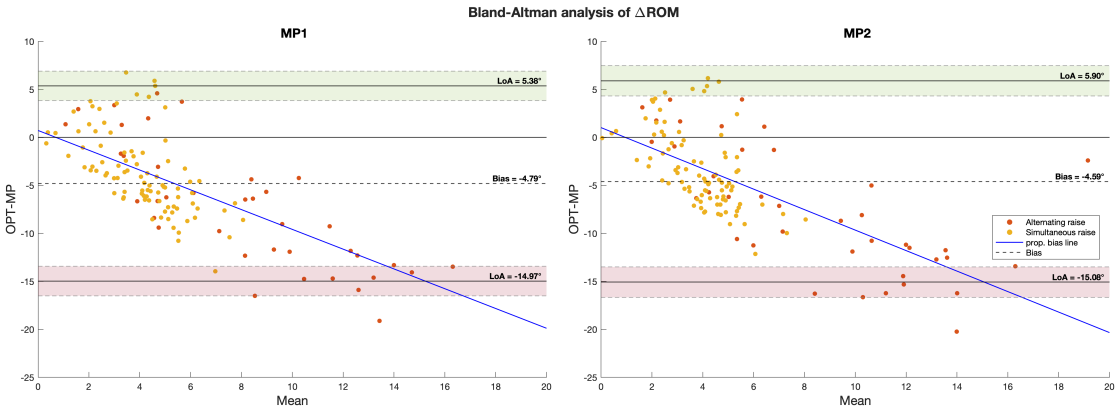

Figure 4.27: Bland-Altman analysis of ΔROM parameter, conducted for MP1 and MP2 during **frontal raises**. Each task is color-coded: red for alternating raises and yellow for simultaneous ones.

Greater agreement between the two systems is evident for parameters characterized by higher mean values, such as ΔRV. In lateral raises (Figure 4.28), the LoA for MP1 delimits an interval of [-10.22°/s, 10.22°/s], with a bias of zero. Agreement improves when the variability derived by the segmentation is removed: in fact, in MP2, the LoA define a narrower interval, between -7.69°/s and -7.05°/s with a slight negative bias of -0.31°/s. A similar tendency is appreciable in frontal raises

(Figure 4.29), where both a variation in the bias and a reduction of the interval delimited by the LoA are observed.

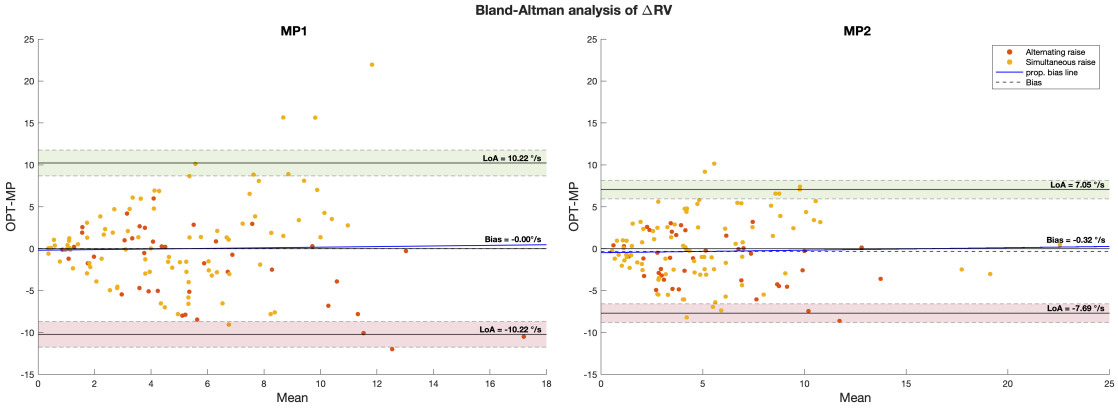

Figure 4.28: Bland-Altman analysis of ΔRV parameter, conducted for MP1 and MP2 during **lateral raises**. Each task is color-coded: red for alternating raises and yellow for simultaneous ones.

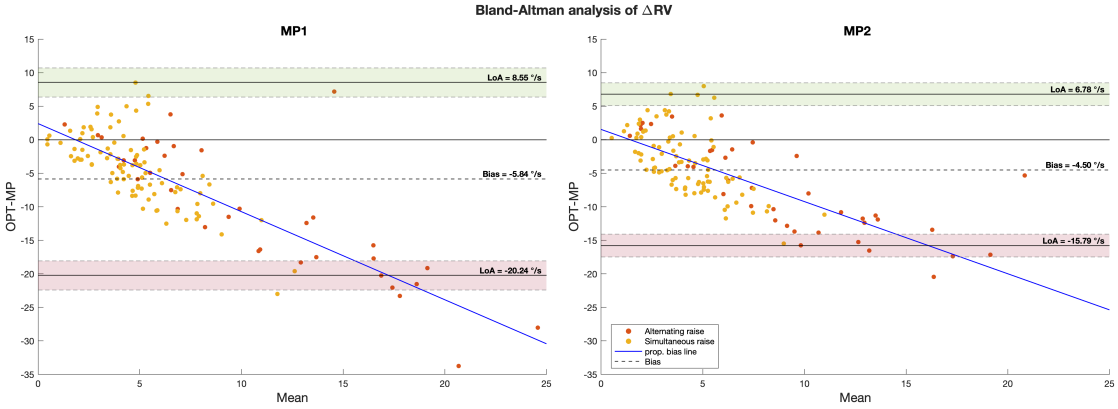

Figure 4.29: Bland-Altman analysis of ΔRV parameter, conducted for MP1 and MP2 during **frontal raises**. Each task is color-coded: red for alternating raises and yellow for simultaneous ones.

In conclusion, for bilateral parameters, the variability induced by the standard segmentation process is significant only for ΔRV. This behavior can be explained since the time intervals used to compute this parameter are no longer influenced by segmentation variability.

4.4 Parameters in parkinsonian subjects

This section reports the average parameters computed for parkinsonian subjects for descriptive purposes. The inclusion of these results aims to demonstrate the feasibility of data acquisition using the exergame, as well as to demonstrate the potential of the developed pipeline to extract angular trajectories and outcome parameters in the final application scenario.

Domain	Parameter	N_{lat}	Lateral (Mean±SD)	N_{front}	Frontal (Mean±SD)
Angular	HA [deg]	67	112.80 ± 16.65	72	138.99 ± 24.35
	LA [deg]	67	25.88 ± 6.66	72	24.51 ± 7.42
	ROM [deg]	67	86.67 ± 18.73	72	114.23 ± 28.56
	OR [deg]	67	62.80 ± 16.65	72	88.99 ± 24.35
Temporal	PPT [s]	67	3.39 ± 1.21	72	3.11 ± 1.18
	PPM [peaks/minute]	67	20.34 ± 7.29	72	22.29 ± 6.51
Velocity-based	RV [deg/s]	67	89.64 ± 23.18	72	115.64 ± 29.03
Spectral	PF [Hz]	67	0.28 ± 0.12	72	0.30 ± 0.15

Table 4.23: Unilateral parameters computed for parkinsonian subjects, expressed as mean±SD, for both lateral and frontal raises. The number of trials used (N_{lat} and N_{front}) is also reported for each condition.

Domain	Parameter	N_{lat}	Lateral (Mean±SD)	N_{front}	Frontal (Mean±SD)
Angular	ΔHA [deg]	22	10.06 ± 7.97	24	10.01 ± 9.12
	ΔLA [deg]	22	2.59 ± 2.17	24	2.59 ± 2.74
	ΔROM	22	9.97 ± 9.07	24	9.96 ± 9.59
Temporal	IPID [s]	22	0.30 ± 0.43	24	0.22 ± 0.43
	PTL [s]	22	1.50 ± 1.72	24	1.39 ± 1.89
	CCL [s]	22	1.07 ± 1.24	24	0.77 ± 1.29
Velocity-based	ΔRV [deg/s]	22	11.66 ± 10.06	24	10.04 ± 10.73
	ΔMV [deg/s]	22	1.33 ± 1.04	24	0.99 ± 1.76

Table 4.24: Bilateral parameters computed for parkinsonian subjects, expressed as mean±SD, for both lateral and frontal raises. The number of trials used (N_{lat} and N_{front}) is also reported for each condition.

Table 4.23 shows the mean values and the respective standard deviations (SD) of the unilateral parameters for both lateral and frontal raises. Some trials were excluded from the analysis, as some subjects were unable to perform the exercise. These exclusions were limited to the lateral raises, which were probably more difficult to perform for this kind of subjects.

Table 4.24 reports the bilateral parameters for both lateral and frontal raises. Even in this case, some lateral tasks were excluded as some subjects were unable to perform the movement.

Each parameter exhibits a high standard deviation: this is to be expected as subjects present different levels of motor impairment, leading to a high inter-subject variability. This aspect suggests that the system is capable to detect significant

differences in the movements of Parkinsonian subjects, a fundamental element for a clinical evaluation.

In the lateral execution, lower mean values are observed compared to those relative to the frontal execution. This may indicate a more pronounced difficulty for the subjects in performing lateral movements. It is important to underline that, as observed in the validation analysis, even in healthy subjects, the parameters related to the frontal raises tend to be higher both in terms of mean value and standard deviation, than those observed in lateral raises.

The bilateral parameters present similar mean values and standard deviations across both executions. These coordination parameters, computed on individuals with asymmetrical motor abilities between the two arms, are capable to effectively capture the deviation between the right and left limbs, contrary to what was observed in the group of healthy subjects.

The analysis of the angular trajectories showed that those related to frontal raises in parkinsonian subjects were more regular than those observed in healthy subjects. This regularity represents an advantage, as it reduces the intrinsic variability of the signal. An example of an angular trajectory from a single-arm frontal raise is shown in Figure 4.30.

Although some trials were performed while the subject was in a seated position, this had no significant impact on the tracking quality (Figure 4.31); despite not all the landmarks were completely tracked.

The relatively clean trajectories obtained even in non-optimal conditions suggest that the quality of the tracking is strongly influenced by movement speed, which is substantially lower in parkinsonian subjects than in healthy subjects.

In conclusion, these findings suggest that the MP model allows accurate assessment of motor performance through the exergame *Palestra* in the target population.

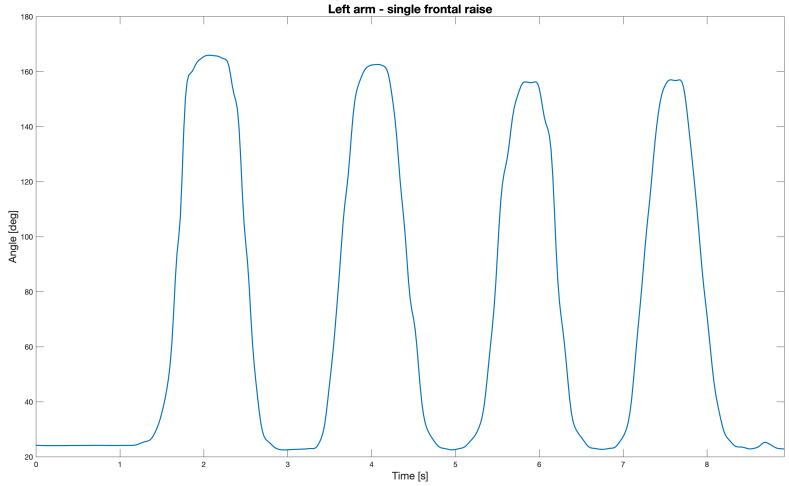

Figure 4.30: Angular trajectory related to a single-arm frontal raise of a parkinsonian subject. The signal exhibits a smoother profile than those observed in healthy subjects.

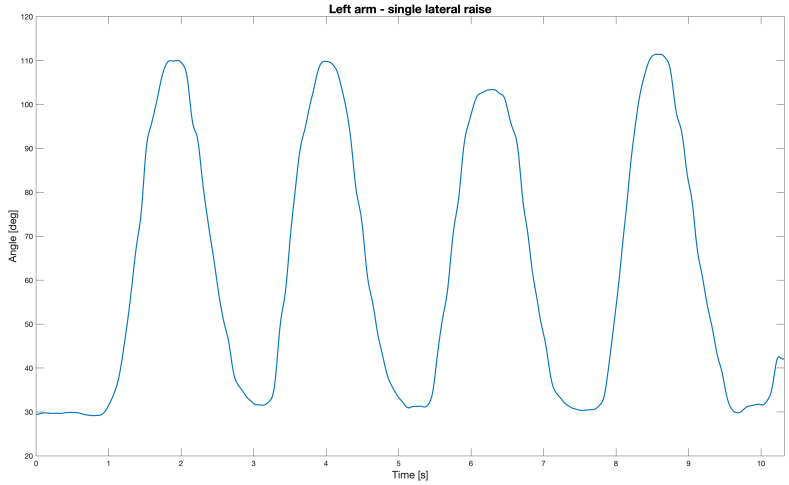

Figure 4.31: Angular trajectory obtained from a subject who performed the task in a seated condition. The signal does not appear to be affected by the non-optimal execution posture.

Chapter 5

Discussions

The exergame was developed as a cost-effective and broadly accessible tool aimed at supporting motor training of the upper limbs. The main objectives is to collect data continuously that allows for remote analysis and provides clinicians with general information regarding skills and changes over time.

In this chapter, the results obtained are discussed in relation to the study's goals.

The main results are analysed, highlighting the strengths and limitations of the proposed methodology, with particular attention on the implications of using the exergame. Moreover, the potential future developments are exposed.

5.1 Validation of angular trajectories

The validation analysis conducted on the angular trajectories confirms the reliability of the exergame, both in terms of temporal and spatial resolution.

Temporal domain

Regarding the temporal resolution, despite exergame operating with a variable frame rate, the average value of approximately 30 fps allows the acquisition of signals characterized by a high temporal concordance with those recorded by OPT reference system.

This result is supported by the accurate realignment between the MP and OPT signals, in which key movement events (such as maxima and minima in the angular trajectories) are closely aligned. This temporal agreement is observed for both lateral and frontal raises, as well as for all task types (single arm, alternating and simultaneous raises). The Bland-Altman analysis confirms this observation, revealing an excellent agreement in all temporal parameters computed. For instance,

the PPT parameter exhibits incredibly narrow Limits of Agreement: [-0.09 s, 0.09 s] in lateral raises and [-0.23 s, 0.24 s] in frontal raises.

A great temporal resolution is essential for the accurate computation of temporal parameters. In fact, even in ideal conditions in which the intrinsic variability introduced by the segmentation was negligible, if the system fails to identify key movement events, at the moment in which they occur, the resulting temporal parameters would not be reliable.

This aspect could be critical for assessing, for example, the patient reactivity associated with cognitive functions. For instance, in "manual" control mode, the subject must perform the movement in response to a visual cue. In such case, the time span between the visualization of the cue and the motor response can be measured, thus obtaining useful information of the cognitive abilities of the subject.

These findings indicates that the system is sufficiently responsive to the user movements, allowing the assessment not only of motor, but also of cognitive functions, which can be further explored in future studies.

Spatial domain

The application of the developed analysis pipeline produced high-quality angular trajectories, in both lateral and frontal raises. As reported in the *Result* section 4.1, the analysis performed on the lateral raises yielded lower error metric values compared to the frontal ones. Furthermore, the alignment between MP and OPT is nearly perfect in the rising and falling edges of the movement. This latter aspect is further confirmed by the high linear correlation coefficient (ρ=0.99).

The trajectories recorded in the lateral raises are more similar to those of the optoelectronic system that those observed in the frontal raises. This morphological similarity may be attributed to the nature of the movement, which occurs parallel to the camera plane. In this condition, the x and y coordinates exhibit larger excursion and are calculated more precisely by Google MediaPipe Pose. In contrast, the z coordinate (depth), which is only estimated by the model, presents minimal variations during lateral raises, and therefore does not significantly affect the computation of the 3D angular trajectories. As the z coordinate is only an estimation, it could introduces errors and variability in the tracking, but in lateral raises the z excursion is very narrowed and hence does not affect the morphology of the signal.

Conversely, in the frontal raises, the movement occurs in a plane perpendicular to the camera, where the z coordinate is crucial for computing angular trajectories. However, since the z coordinate is only estimated, it could introduce random fluctuations in the movement trajectories, in particular at transition points, (i.e., the transition from minima to maxima and vice versa), especially when those occur rapidly.

This behaviour emerged from the analysis of the z coordinate in the frontal raises (Figure 5.1), where greater instability was observed at transition points, compared to the y coordinate. This phenomenon is reflected in the morphology of the angular trajectories, which appear more "squared" and present more marked random fluctuations, both in the maximum and minimum points.

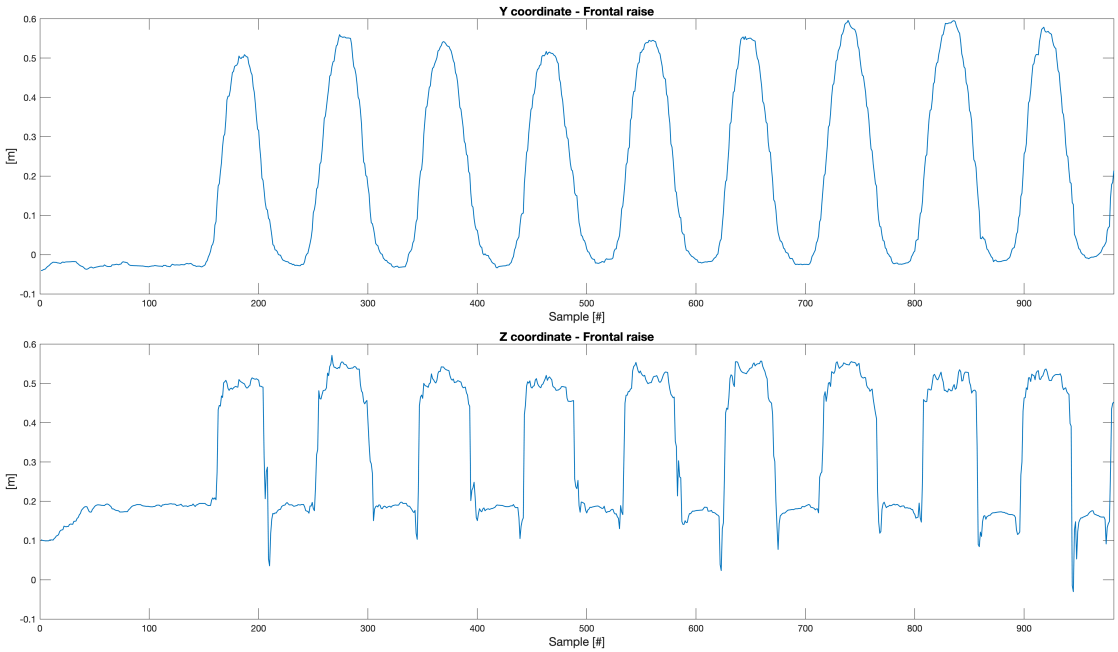

Figure 5.1: Unprocessed y and z coordinates during frontal raises. The z signal is characterized by spikes and random fluctuation in transition points while the y profile remains smoother.

The analysis of the individual tasks indicates that the single-arm and simultaneous raises present trajectories more similar to the reference, in both lateral and frontal raises, compared to alternating raises. The latter are affected by the transient undulation problem described in *chapter* 3 in *section* 3.2, probably caused by Google MediaPipe Pose, that tends to generate a sort of "bounce" in the landmarks associated with the stationary limb while the opposite limb is in motion.

This effect manifests itself as a concave wave in lateral raises, and as a convex wave in frontal raises. In the latter case, this artifact paradoxically results in a lower RMSE than the other frontal tasks, as the transient occurs in signal intervals where the differences between the MP and OPT signals are lower. In fact, in frontal executions, the RMSE for alternating raises is $9.755 \pm 2.434°$, compared to $10.145 \pm 3.458°$ for single-arm raises and $12.516 \pm 2.747°$ for simultaneous raises.

In both lateral and frontal raises, even in all the tasks analyzed, an overestimation of angular values by the MP system with respect to OPT is observed, both for the minimum and maximum points of the trajectories.

The Bland-Altman analysis of the LA and HA parameters, computing using the OPT-based segmentation, reveals systematic bias that differs between minimum and maximum points. For the minimum points (LA), MP shows a bias of -7.55° in lateral raises and -12.34° in frontal raises. Regarding the maximum points (HA), the bias is -4.08° in lateral raises and -12.34° in frontal raises. This behaviour could be due to differences between the anatomical landmark positions estimated by MP model and those obtained by the OPT marker placement. These differences could introduce variability that reflects as overestimation in the angular values computed through MP.

This overestimation was not observed in two of the subjects enrolled in the study, who showed trajectories nearly perfect aligned between MP and OPT. This could be explained by various factors, including the anthropometric measurements of the subject, the clothing worn during data acquisition, the speed of movement and, in particular, by a better marker placement in OPT, which is presumably the most influential factor.

Furthermore, the Bland-Altman analysis suggests that these discrepancies are more dependent by subject-specific factors than by the type of task performed. In fact, in the two subjects previously mentioned, the trajectory alignment was excellent across all tasks performed. This is clearly visible in the lateral raises, and also pronounced in the frontal raises, where only minor discrepancies were observed. This behaviour suggests the presence of a relationship between the subject morphology and the estimation of the MP model, which will have to further investigate in future studies.

A key point to take in account is that the processed angular trajectories themselves could provide clinically relevant information for the qualitative description of the movement. As confirmed by the validation analysis, the morphology of the trajectories, except for some random fluctuations easily identifiable upon visual inspection, follows the same morphology as the reference trajectories. By analysing these trajectories, it is possible to deduce relevant information such as, for example, the subject's level of fatigue, movements slowdowns, or difficulties encountered during the executions of certain tasks. All this information is fundamental for the clinician's decision-making process, and the possibility of obtaining such data remotely, continuously and in real time allows saving time and resources, enhancing the overall effectiveness of the patient monitoring.

In summary, the results confirm the practical relevance and utility of the exergame and the body tracking using Google MediaPipe Pose in the context of telemonitoring and telerehabilitation. In such settings, the high level of accuracy provided by optoelectronic system is not required. Rather, it is essential to have reliable data, that allows continuous patient monitoring and supports the clinician throughout the rehabilitation process.

5.2 Agreement between systems and acceptability of the results in the application context

The validation analysis performed on both the mean parameters and those computed from the individual repetitions, allows for the evaluation of the tracking accuracy. It allowed for the identification of parameters that are suitable for use within the intended application context, as well as those that show critical issues, considering lateral and frontal executions separately.

Unilateral parameters

Both from the analysis of the aggregate parameters and that conducted on single repetitions indicate that, in general, MP tends to overestimate most angular parameters compared to the optoelectronic system. This behaviour is consistent with what was noted from the angular trajectories, where non-proportional biases were evident, both in the maxima and minima points, across both lateral and frontal raises.

Overall, parameters related to the lateral raises exhibit lower error metrics and standard deviations compared to those computed during frontal executions. This trend confirms what was already found in the trajectory analysis, where greater morphological variability in the angular trajectories of frontal raises adversely influences the computation of the parameters.

The Bland-Altman analysis performed on the single repetitions enabled an evaluation of the agreement between the two systems, highlighting the presence of any systematic bias and the relative error associated with each parameter.

The LA parameter, which is influenced by both the systematic bias and the segmentation algorithm, exhibits relatively wide Limits of Agreement (LoA). In lateral raises, the LoA define an interval ranged from -9.16° to 13.22°, while in frontal raises, they delimit a wider interval of [-23.53°, 16.42°]. These relatively wide intervals could be attributed to differences between the position of the landmarks defined by the MP model and those determined by the marker placement in OPT. In the frontal raises the band of agreement is particularly wide, limiting the applicability of the LA parameter for motor assessment, even in an approximate manner, within the intended application context.

This limited reliability is further supported by the presence of a limited negative correlation, which make the interpretation of the LA parameter more difficult, even if supported by visual analysis of the trajectories.

In contrast, ROM shows acceptable agreement interval within the intended application context, in particular in the lateral raises, where the interval ranges from [-18.16°, 3.29°]. Furthermore, it exhibits a strong correlation in lateral raises, suggesting that the two measures vary coherently (if one measure increases, the

other also increases). This consistency allows a reliable evaluation of the parameter, especially when accompanied by visual analysis of the corresponding angular trajectory.

Even in the presence of an error equal to -18.16°, if the entire signal is characterized by the presence of a systematic bias, the average value of the parameter may still be used as a valid indicator for motor assessment. Indeed, all mean parameters computed exhibit a high level of agreement between MP and OPT, both in terms of mean and variability. This tendency is particularly evident in lateral raises, although still present in frontal raises. In general, this indicates the reliability of mean parameters as quantitative metrics of the subject's motor performance.

The velocity parameter RV is also acceptable, for the same considerations above.

All parameters computed over the entire signal, thus not directly influenced by segmentation (except for PPM, computed through segmentation), exhibit a notable reliability. The average values obtained for these parameters are practically identical between the two systems, in both lateral and frontal raises, confirming the suitability of the exergame for use in rehabilitation setting and telemonitoring.

The PPT temporal parameter reveals an almost perfect level of agreement between the two systems, both in the lateral and frontal raises. This result is supported by practically identical mean values, strong correlation coefficient, and by the extremely narrow limits of agreement observed in the Bland-Altman (i.e., [-0.09 s, 0.09 s] in lateral raises, and [-0.23 s, 0.24 s] in frontal raises).

Furthermore, this parameter is sensitive to distinctions between different task types. In particular, the Bland-Altman analysis reveals a separate distribution in the measurements related to the alternating raises, forming a well-defined cluster compared to the those of single-arm and simultaneous tasks. This clusterization is consistent with the physiologically higher time intervals between two consecutive peaks, which is a characteristic of the alternating movements.

These findings confirm the excellent reliability of temporal parameters, which prove to be an excellent metric for an accurate motor assessment.

In general, it should be noted that each parameter must be interpreted in relation to the trajectory from which it is derived. The analysis of the trajectories has revealed that the signal generated by the exergame accurately follows the pattern of the movement, thus following the reference signal's morphology. Consequently, even if a parameter presents an apparently high error, it may still provide relevant information when evaluated together with the visual analysis of the trajectory. In this case, parameters are not intended to replace the information obtained through the trajectory observation, instead to integrate it with quantitative information.

In conclusion, all unilateral parameters, except for LA, can be considered valid metrics to support the analysis of the movement, providing a quantitative complement to the visual assessment of angular trajectories.

Bilateral parameters

All bilateral parameters, whether aggregated or extracted from individual repetitions, show comparable values between the two systems. This result is expected, since all participants enrolled in the study were healthy and presented perfect symmetry in limb movements. The correlation analysis highlighted weak or moderate correlations for most parameters, other than temporal ones, confirming the high reliability of temporal measurement. The weak correlation observed in the remaining parameters, especially in the angular ones, may be attributable to the nature of their computation, as they are defined as difference between the values of each limb. In the presence of nearly perfect symmetry, such differences do not show informative movement information but rather represent the residual variability of the signals. This is further confirmed by the Bland-Altman analysis, which shows a random dispersion in the measurements.

For a meaningful assessment of these bilateral parameters, it would be necessary to acquire data from subjects with motor asymmetries, to evaluate the system's capability to capture inter-limb differences.

Similarly, it can also be confirmed that the temporal parameters are reliable and less susceptible to variability than parameters in other measurement domains, in both execution modes.

Final considerations

In conclusion, all temporal parameters, both unilateral and bilateral, proved to be reliable and very accurate, in lateral raises as well as in frontal ones. A comparable consistency was observed for parameters computed over the entire signal, such as PPM, CCL and PF, which confirm their validity as reliable measures in the intended application context of the exergame. Regarding angular parameters, with the exception of LA, it can be stated that they are sufficiently accurate for measuring and evaluating upper limb motor functions through the exergame and the MP model. However, their interpretation should always be integrated by visual analysis of the angular trajectories, to offer further qualitative information for enhancing the overall assessment.

Finally, it is important to consider that the accuracy of the MP tracking model is usually lower for frontal raises than lateral raises, and it should be considered for the interpretation of the outcomes.

5.3 On the effects of the segmentation algorithm

The developed pipeline proved to be reliable, robust and able to produce high quality angular trajectories suitable for motor assessment.

One of the key elements of the pipeline is the application of the moving median filter, which demonstrated to be effective in reducing spikes and random fluctuations, especially in trajectories related to the z coordinate.

The segmentation algorithm appeared to be robust and accurate, even in morphologically variable trajectories, such as those associated with frontal raises. Misidentification of maxima and minima (false positive) were rare, suggesting stable behaviour even in the presence of non-ideal signals. These potential errors inevitably introduce additional variability, which may compromise the accuracy of parameters.

This variability was observed through the validation analysis based on the segmentation derived from the OPT signals. Except for the LA parameter and bilateral outcomes, the Bland-Altman analysis revealed that the limits of agreement (LoA) narrowed in all parameters, confirming that segmentation introduces a source of variability in the computation of the parameters.

However, the improvements observed were not sufficiently substantial to change the overall levels of agreement between the two systems. This result confirms that, whereas the developed segmentation algorithm introduces a certain grade of variability, it still guarantees an appropriate level of reliability for the intended use of the exergame.

5.4 Applicability on subjects with Parkinson's disease

The exergame was tested on 12 parkinsonian subjects, characterized by different levels of motor impairment. The testing was carried out during the initial phase of the ELEVATOR (NODES PNRR program) project experimentation, for which the exergame was conceptualized and developed by CNR-IEIIT.

The system did not show any particular tracking issue towards the target population, even under potentially critical conditions, such as tasks performed in a seated position. The angular trajectories obtained were generally analyzable but some trials were excluded because the subject was unable to perform the proposed exercise. Overall, it can be stated that almost all the subjects enrolled were able to complete the proposed tasks, except of those exhibiting severe motor disabilities.

In the absence of a control group and a reference measurement system, no proper statistical analysis could be carried out. The results obtained should therefore be interpreted as a first assessment of the applicability of the system.

The angular trajectories related to the frontal raises were unexpectedly characterized by lower variability compared to those obtained in healthy subjects. This behaviour could be due to the reduced execution speed, which seems to ensure tracking robustness and mitigate the effects of tracking fluctuations.

The extracted parameters showed to be consistent with the visual inspection of the related angular trajectories, confirming their ability to provide complementary and relevant quantitative information.

In summary, the test conducted on the target population highlighted the following key points:

- The system did not show any limitations neither in tracking slow unstable movements (e.g., freezing arm lifts) nor movements performed in partial occlusion conditions (i.e., seated posture).
- The enrolled subjects completed the proposed tasks without difficulty, except for those characterized by a high level of motor impairment.
- The pipeline generated high quality angular trajectories, appropriate for motor performance analysis.
- The parameters were consistent with the observation derived from visually inspecting the trajectories.

In conclusion, this first experimental test indicates that the exergame produced promising results, supporting its potential application in a telerehabilitation context.

5.5 Limits and future developments

The study conducted has provided positive and encouraging findings, confirming the usability of the system in the designated application scenario. Nevertheless, there are some limitations that should be considered to properly understand the findings and support future research.

The main limitation concerns the composition of the sample included for the validation analysis, which did not include Parkinsonian subjects. Unfortunately, due to logistical and administrative barriers, it was not possible to enroll these subjects to carry out the acquisitions in the laboratory. The inclusion of such data would have outlined both the prospects and the weaknesses of the system even in this condition, reinforcing the overall validation process. The inclusion of data from parkinsonian subject would have allowed for a deeper evaluation of the bilateral parameters, which are designed for the specific purpose of detecting asymmetries in upper limb movements.

Another limitation regards the lack of data involving movements performed at different speeds. To deepen the tracking stability of the exergame under different conditions, it would have been appropriate to include acquisitions representing at least 3 different speed levels (i.e., slow, moderate and fast).

An important constraint, which currently limits the applicability of the system, concerns the dependence on MATLAB for data processing. This makes it difficult

to produce a complete report, with angular trajectories and outcome parameters, for users who do not own a license for this tool.

Another relevant aspect that could not be explored concerns the influence of the environmental conditions on tracking stability. For example, performing the proposed tasks under different lighting conditions, would enable to establish guidelines for ensuring a better user experience.

Furthermore, possible errors in the marker placement process may have slightly influenced the validation results. These could be due to the variability present in the placement on each subject. These errors, albeit small, should be considered during result interpretations.

Considering these limitations, the following future developments are proposed:

- Include in the experimental protocol trials characterized by movement with different levels of speed and asymmetric raises, acquired under different environmental conditions. This allows for the evaluation of the system's robustness in more challenging scenarios.
- Integrate the entire developed pipeline into the exergame, to display a complete report directly within the system.
- Analyze the impact of markerization for the optoelectronic system, to correct, if necessary, small inaccuracies in the comparison between systems.

This work represents an initial stage of a broader project. In the future, it will be necessary to collaborate with neurologists to evaluate in detail the applicability of the system in clinical practice, and possibly implement additional features, such as the computation of new parameters, to respond to new clinical needs that have not yet been addressed.

Chapter 6

Conclusions

In this work, kinematic data were acquired using the exergame *Palestra*, subsequently analyzed through a specifically developed pipeline, to extract outcome parameters for a preliminary quantitative motor assessment. A validation procedure was conducted using data collected from five healthy subjects, to evaluate the reliability and accuracy of both the exergame and the processing pipeline; for this purpose, a total of 80 trajectories was analyzed. Finally, *Palestra* was tested on a target user group, consisting of parkinsonian patients characterized by different levels of motor impairment.

The results obtained confirm the attainment of the established goals:

- The developed pipeline proved to be robust and reliable, producing good quality angular trajectories showing an excellent temporal resolution.
- The validation analysis demonstrated good agreement between the MP model and the reference optoelectronic system, especially in the lateral raises. Angular parameters proved to be sufficiently accurate for the telemonitoring context, while temporal parameters exhibited a high reliability. Trajectories and parameters related to the frontal executions also proved to be acceptable for the required application context, although they present a lower accuracy than lateral raises.
- The system was also tested without any technical issues on parkinsonian subjects, obtaining promising preliminary results in terms of applicability in clinical setting.

In the future, several challenges should be resolved, such as the extension of the system in clinical practice, the integration of new functionalities to meet emerging clinical needs, and resolving the methodological constraints identified in the present study.

Innovation in the field of telemedicine, especially in telerehabilitation, represents a fundamental aspect to explore and deepen in the near future, as it is one of the sectors that meets the rising demand to enhance the Quality of Life (QoA) in an ageing population. Tools such as *Palestra* allow the opportunity to perform rehabilitation tasks at home in a continuous, economical, and engaging manner. Motor performance data can be collected continuously and analyzed remotely, enabling the referring clinician to perform preliminary and timely assessment of the patient's motor progress. A system like this can significantly improve the quality of care provided to the patient, simplifying the clinical decision-making process, and reducing the number of visits to congested healthcare facilities; therefore, it supports the implementation of personalized treatment plans.

In conclusion, *Palestra* represents a practical, economical, and accessible solution for upper limb motor training and recovery. The system offers new opportunities in the field of telemedicine and home-based rehabilitation, positioning itself as a reliable tool, capable of generating high-quality data with minimal costs. Therefore, *Palestra* has the potential to improve the quality of life for elderly subjects and for those affected by neurodegenerative diseases, supporting clinician in the daily management of these conditions.

Bibliography

[1] World Health Organization. *World Report on Ageing and Health*. "Geneva, Switzerland": World Health Organization, 2015 (cit. on p. 1).

[2] Luca Vismara, Claudia Ferraris, Gianluca Amprimo, Giuseppe Pettiti, Francesca Buffone, Andrea Gianmaria Tarantino, Alessandro Mauro, and Lorenzo Priano. «Exergames as a rehabilitation tool to enhance the upper limbs functionality and performance in chronic stroke survivors: a preliminary study». In: *Frontiers in Neurology* 15 (2024) (cit. on pp. 1, 2).

[3] Yoonsin Oh and Stephen Yang. «Defining exergames & exergaming». In: *Proceedings of meaningful play* 2010 (2010), pp. 21–23 (cit. on p. 1).

[4] Marc A Adams, Simon J Marshall, Lindsay Dillon, Susan Caparosa, Ernesto Ramirez, Justin Phillips, and Greg J Norman. «A theory-based framework for evaluating exergames as persuasive technology». In: *Proceedings of the 4th international conference on persuasive technology*. 2009, pp. 1–8 (cit. on p. 1).

[5] Marianne Hernholm. «A Virtual Reality Pose Estimation Exercise Game for Post-Stroke Upper-Limb Motor Function Rehabilitation». Master's Thesis. MA thesis. Norway: Norwegian University of Science and Technology, June 2023. URL: `https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3095629?show=full` (cit. on p. 1).

[6] Reza Haghighi Osgouei, David Soulsby, Fernando Bello, et al. «Rehabilitation exergames: Use of motion sensing and machine learning to quantify exercise performance in healthy volunteers». In: *JMIR Rehabilitation and Assistive Technologies* 7.2 (2020), e17289 (cit. on p. 1).

[7] Papamichael Elena, Solou Demetris, Michailidou Christina, and Papamichail Marios. «Differences between exergaming rehabilitation and conventional physiotherapy on quality of life in Parkinson's disease: a systematic review and meta-analysis». In: *Frontiers in neurology* 12 (2021) (cit. on p. 1).

[8] Mauro Callejas-Cuervo, Gloria M Díaz, and Andrés Felipe Ruíz-Olaya. «Integration of emerging motion capture technologies and videogames for human upper-limb telerehabilitation: A systematic review». In: *Dyna* 82.189 (2015), pp. 68–75 (cit. on p. 2).

[9] Gioia Mura, Mauro G Carta, Federica Sancassiani, Sergio Machado, Luca Prosperini, et al. «Active exergames to improve cognitive functioning in neurological disabilities: a systematic review and meta-analysis». In: *European journal of physical and rehabilitation medicine* 54.3 (2018), pp. 450–462 (cit. on p. 2).

[10] João Abreu, Sérgio Rebelo, Hugo Paredes, João Barroso, Paulo Martins, Arsénio Reis, Eurico Vasco Amorim, and Vítor Filipe. «Assessment of microsoft kinect in the monitoring and rehabilitation of stroke patients». In: *Recent Advances in Information Systems and Technologies: Volume 2 5*. Springer. 2017, pp. 167–174 (cit. on p. 3).

[11] Justin Amadeus Albert, Victor Owolabi, Arnd Gebel, Clemens Markus Brahms, Urs Granacher, and Bert Arnrich. «Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: A pilot study». In: *Sensors* 20.18 (2020), p. 5104 (cit. on pp. 3, 10).

[12] Burakhan Çubukçu, Uğur Yüzgeç, Raif Zileli, and Ahu Zileli. «Reliability and validity analyzes of Kinect V2 based measurement system for shoulder motions». In: *Medical engineering & physics* 76 (2020), pp. 20–31 (cit. on p. 3).

[13] Google. *Pose Landmarker*. 2024. URL: `https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker?hl=it` (cit. on p. 3).

[14] Dale M Harris, Timo Rantalainen, Makii Muthalib, Liam Johnson, and Wei-Peng Teo. «Exergaming as a viable therapeutic tool to improve static and dynamic balance among older adults and people with idiopathic Parkinson's disease: a systematic review and meta-analysis». In: *Frontiers in aging neuroscience* 7 (2015), p. 167 (cit. on p. 5).

[15] Błażej Cieślik, Justyna Mazurek, Adam Wrzeciono, Lorenza Maistrello, Joanna Szczepańska-Gieracha, Pierfranco Conte, and Pawel Kiper. «Examining technology-assisted rehabilitation for older adults' functional mobility: a network meta-analysis on efficacy and acceptability». In: *Npj Digital Medicine* 6.1 (2023), p. 159 (cit. on p. 6).

[16] Augusto Garcia-Agundez, Ann-Kristin Folkerts, Robert Konrad, Polona Caserman, Thomas Tregel, Mareike Goosses, Stefan Göbel, and Elke Kalbe. «Recent advances in rehabilitation for Parkinson's Disease with Exergames: A Systematic Review». In: *Journal of neuroengineering and rehabilitation* 16 (2019), pp. 1–17 (cit. on p. 6).

[17] Dijana Nuic, Maria Vinti, Carine Karachi, Pierre Foulon, Angèle Van Hamme, and Marie-Laure Welter. «The feasibility and positive effects of a customised videogame rehabilitation programme for freezing of gait and falls in Parkinson's disease patients: a pilot study». In: *Journal of neuroengineering and rehabilitation* 15 (2018), pp. 1–11 (cit. on pp. 6, 11).

[18] Ioannis Pachoulakis, Nikolaos Papadopoulos, and Anastasia Analyti. «Kinect-Based Exergames Tailored to Parkinson Patients». In: *International Journal of Computer Games Technology* 2018.1 (2018), p. 2618271 (cit. on p. 6).

[19] Gianluca Amprimo, Giulia Masi, Lorenzo Priano, Corrado Azzaro, Federica Galli, Giuseppe Pettiti, Alessandro Mauro, and Claudia Ferraris. «Assessment tasks and virtual exergames for remote monitoring of Parkinson's disease: An integrated approach based on Azure Kinect». In: *Sensors* 22.21 (2022), p. 8173 (cit. on p. 6).

[20] Claudia Ferraris et al. «Usability of the REHOME solution for the telerehabilitation in neurological diseases: Preliminary results on motor and cognitive platforms». In: *Sensors* 22.23 (2022), p. 9467 (cit. on p. 7).

[21] Nahid Norouzi-Gheidari, Alejandro Hernandez, Philippe S Archambault, Johanne Higgins, Lise Poissant, and Dahlia Kairy. «Feasibility, safety and efficacy of a virtual reality exergame system to supplement upper extremity rehabilitation post-stroke: a pilot randomized clinical trial and proof of principle». In: *International journal of environmental research and public health* 17.1 (2020), p. 113 (cit. on pp. 7, 11).

[22] Elise Klæbo Vonstad, Xiaomeng Su, Beatrix Vereijken, Kerstin Bach, and Jan Harald Nilsen. «Comparison of a deep learning-based pose estimation system to marker-based and kinect systems in exergaming for balance training». In: *Sensors* 20.23 (2020), p. 6940 (cit. on p. 7).

[23] Andrea Avogaro, Federico Cunico, Bodo Rosenhahn, and Francesco Setti. «Markerless human pose estimation for biomedical applications: a survey». In: *Frontiers in Computer Science* 5 (2023), p. 1153160 (cit. on pp. 7, 9).

[24] Huaiming Ji, Li Wang, Yuwei Zhang, Zhi Li, and Chenglong Wei. «A review of human pose estimation methods in markerless motion capture». In: *Computer-Aided Design and Applications* (2023), pp. 392–423 (cit. on p. 8).

[25] Thomas Hellsten, Jonny Karlsson, Muhammed Shamsuzzaman, and Göran Pulkkis. «The potential of computer vision-based marker-less human motion analysis for rehabilitation». In: *Rehabilitation Process and Outcome* 10 (2021), p. 11795727211022330 (cit. on pp. 8, 11).

[26] Matthew Pardell, Naomi D Dolgoy, Stéphanie Bernard, Kerry Bayless, Robert Hirsche, Liz Dennett, and Puneeta Tandon. «Movement Outcomes Acquired via Markerless Motion Capture Systems Compared with Marker-Based Systems for Adult Patient Populations: A Scoping Review». In: *Biomechanics* 4.4 (2024), pp. 618–632 (cit. on p. 8).

[27] Kyle L Jackson, Zoran Durić, Susannah M Engdahl, Anthony C Santago II, Secili DeStefano, and Lynn H Gerber. «Computer-assisted approaches for measuring, segmenting, and analyzing functional upper extremity movement: a narrative review of the current state, limitations, and future directions». In: *Frontiers in Rehabilitation Sciences* 4 (2023), p. 1130847 (cit. on p. 8).

[28] Gianluca Amprimo, Giulia Masi, Giuseppe Pettiti, Gabriella Olmo, Lorenzo Priano, and Claudia Ferraris. «Hand tracking for clinical applications: validation of the Google MediaPipe Hand (GMH) and the depth-enhanced GMH-D frameworks». In: *Biomedical Signal Processing and Control* 96 (2024), p. 106508 (cit. on pp. 8, 14).

[29] Enrico Martini, Michele Boldo, Stefano Aldegheri, Nicola Valè, Mirko Filippetti, Nicola Smania, Matteo Bertucco, Alessandro Picelli, and Nicola Bombieri. «Enabling gait analysis in the telemedicine practice through portable and accurate 3D human pose estimation». In: *Computer Methods and Programs in Biomedicine* 225 (2022), p. 107016 (cit. on p. 8).

[30] Bradley Scott, Martin Seyres, Fraser Philp, Edward K Chadwick, and Dimitra Blana. «Healthcare applications of single camera markerless motion capture: a scoping review». In: *PeerJ* 10 (2022), e13517 (cit. on pp. 9, 11).

[31] Veronica Cimolin et al. «Computation of gait parameters in post stroke and Parkinson's disease: a comparative study using RGB-D sensors and optoelectronic systems». In: *Sensors* 22.3 (2022), p. 824 (cit. on p. 9).

[32] Logan Wade, Laurie Needham, Polly McGuigan, and James Bilzon. «Applications and limitations of current markerless motion capture methods for clinical gait biomechanics». In: *PeerJ* 10 (2022), e12995 (cit. on p. 9).

[33] Claudia Ferraris, Gianluca Amprimo, Giulia Masi, Luca Vismara, Riccardo Cremascoli, Serena Sinagra, Giuseppe Pettiti, Alessandro Mauro, and Lorenzo Priano. «Evaluation of arm swing features and asymmetry during gait in Parkinson's disease using the azure kinect sensor». In: *Sensors* 22.16 (2022), p. 6282 (cit. on pp. 10, 11).

[34] Jamie Shotton et al. «Efficient human pose estimation from single depth images». In: *IEEE transactions on pattern analysis and machine intelligence* 35.12 (2012), pp. 2821–2840 (cit. on p. 10).

[35] Zicheng Liu. *Azure Kinect Body Tracking SDK*. Tech. rep. Microsoft Research, 2020. URL: https://www.microsoft.com/en-us/research/uploads/prod/2020/01/AKBTSDK.pdf (cit. on p. 11).

[36] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. «OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields». In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019) (cit. on p. 11).

[37] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. «Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields». In: *CVPR*. 2017 (cit. on p. 11).

[38] Hao-Shu Fang, Jiefeng Li, Hongyang Tang, Chao Xu, Haoyi Zhu, Yuliang Xiu, Yong-Lu Li, and Cewu Lu. «Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time». In: *IEEE transactions on pattern analysis and machine intelligence* 45.6 (2022), pp. 7157–7173 (cit. on p. 12).

[39] Federico Roggio, Bruno Trovato, Martina Sortino, and Giuseppe Musumeci. «A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review». In: *Heliyon* (2024) (cit. on p. 12).

[40] Marion Mundt, Zachery Born, Molly Goldacre, and Jacqueline Alderson. «Estimating ground reaction forces from two-dimensional pose data: a biomechanics-based comparison of alphapose, blazepose, and openpose». In: *Sensors* 23.1 (2022), p. 78 (cit. on p. 12).

[41] Gaurvi Goyal, Franco Di Pietro, Nicolo Carissimi, Arren Glover, and Chiara Bartolozzi. «Moveenet: online high-frequency human pose estimation with an event camera». In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 4024–4033 (cit. on p. 12).

[42] Ronny Votel and Na Li. *Next-Generation Pose Detection with MoveNet and TensorFlow.js*. Ultimo accesso: 15 maggio 2025. 2021. URL: https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html (cit. on p. 12).

[43] TensorFlow Authors. *MoveNet: Ultra fast and accurate pose detection model*. Ultimo accesso: 15 maggio 2025. 2021. URL: https://www.tensorflow.org/hub/tutorials/movenet (cit. on p. 12).

[44] Jen-Li Chung, Lee-Yeng Ong, and Meng-Chew Leow. «Comparative analysis of skeleton-based human pose estimation». In: *Future Internet* 14.12 (2022), p. 380 (cit. on pp. 12, 13).

[45] BeomJun Jo and SeongKi Kim. «Comparative analysis of OpenPose, PoseNet, and MoveNet models for pose estimation in mobile devices». In: *Traitement du Signal* 39.1 (2022), p. 119 (cit. on p. 12).

[46] Camillo Lugaresi et al. «Mediapipe: A framework for building perception pipelines». In: *arXiv preprint arXiv:1906.08172* (2019) (cit. on p. 13).

[47] Google AI Edge. *MediaPipe: Cross-platform, customizable ML solutions for live and streaming media*. Ultimo accesso: 15 maggio 2025. 2025. URL: `https://github.com/google-ai-edge/mediapipe` (cit. on p. 13).

[48] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. «Blazepose: On-device real-time body pose tracking». In: *arXiv preprint arXiv:2006.10204* (2020) (cit. on p. 13).

[49] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fangfang Zhang, and Matthias Grundmann. *On-Device Real-Time Body Pose Tracking with MediaPipe BlazePose*. Ultimo accesso: 15 maggio 2025. 2020. URL: `https://research.google/blog/on-device-real-time-body-pose-tracking-with-mediapipe-blazepose/` (cit. on p. 13).

[50] Hongyi Xu, Eduard Gabriel Bazavan, Andrei Zanfir, William T Freeman, Rahul Sukthankar, and Cristian Sminchisescu. «Ghum & ghuml: Generative 3d human shape and articulated pose models». In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 6184–6193 (cit. on p. 13).

[51] Valentin Bazarevsky, Ivan Grishchenko, and Eduard Gabriel Bazavan. *Model Card BlazePose GHUM 3D*. Tech. rep. Ultimo accesso: 15 maggio 2025. Google, 2021. URL: `https://storage.googleapis.com/mediapipe-assets/Model%20Card%20BlazePose%20GHUM%203D.pdf` (cit. on p. 13).

[52] Google AI Edge. *MediaPipe Pose Solution*. Ultimo accesso: 15 maggio 2025. 2024. URL: `https://github.com/google-ai-edge/mediapipe/blob/master/docs/solutions/pose.md` (cit. on pp. 13, 14).

[53] Federico Roggio, Sarah Di Grande, Salvatore Cavalieri, Deborah Falla, and Giuseppe Musumeci. «Biomechanical posture analysis in healthy adults with machine learning: Applicability and reliability». In: *Sensors* 24.9 (2024), p. 2929 (cit. on p. 13).

[54] Chang Soon Tony Hii, Kok Beng Gan, Nasharuddin Zainal, Norlinah Mohamed Ibrahim, Shahrul Azmin, Siti Hajar Mat Desa, Bart van de Warrenburg, and Huay Woon You. «Automated gait analysis based on a marker-free pose estimation model». In: *Sensors* 23.14 (2023), p. 6489 (cit. on p. 13).

[55] Sebastian Dill, Andreas Rösch, Maurice Rohr, Gökhan Güney, Luisa De Witte, Elias Schwartz, and Christoph Hoog Antink. «Accuracy evaluation of 3d pose estimation with mediapipe pose for physical exercises». In: *Current Directions in Biomedical Engineering*. Vol. 9. 1. De Gruyter. 2023, pp. 563–566 (cit. on p. 15).

[56] Thiago Buarque De Gusmao Lafayette, Victor Hugo De Lima Kunst, Pedro Vanderlei De Sousa Melo, Paulo De Oliveira Guedes, João Marcelo Xavier Natário Teixeira, Cínthia Rodrigues de Vasconcelos, Veronica Teichrieb, and Alana Elza Fontes Da Gama. «Validation of angle estimation based on body tracking data from RGB-D and RGB cameras for biomechanical assessment». In: *Sensors* 23.1 (2022), p. 3 (cit. on p. 15).

[57] Wouter Durnez, Aleksandra Zheleva, Mark Claypool, Mathias Maes, Klaas Bombeke, Jan Van Looy, and Lieven De Marez. «Spaz! The Effects of Local Latency on Player Actions in a Desktop-Based Exergame». In: *IEEE Transactions on Games* 14.4 (2021), pp. 623–631 (cit. on p. 17).

[58] Laura-Bianca Bilius, Ovidiu Andrei Schipor, and Radu-Daniel Vatavu. «The Age-Reward Perspective: A Systematic Review of Reward Mechanisms in Serious Games for Older People». In: *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences*. 2024, pp. 168–181 (cit. on p. 22).

[59] Qianqian Fang. *JSONLab: a toolbox to encode/decode JSON files in MATLAB/Octave*. `https://iso2mesh.sourceforge.net/cgi-bin/index.cgi?jsonlab`. Accessed: 2025-05-19 (cit. on p. 26).

[60] MathWorks. *findpeaks - MATLAB Function (Signal Processing Toolbox)*. `https://it.mathworks.com/help/signal/ref/findpeaks.html`. Accessed: 2025-05-22 (cit. on p. 33).

[61] MathWorks. *pwelch - Welch's Power Spectral Density Estimate*. `https://it.mathworks.com/help/signal/ref/pwelch.html`. Accessed: 2025-05-19 (cit. on p. 41).

[62] The jamovi project. *jamovi (Version 2.6) [Computer Software]*. Accessed: 2025-05-19. 2024. URL: `https://www.jamovi.org` (cit. on p. 47).

[63] J Martin Bland and DouglasG Altman. «Statistical methods for assessing agreement between two methods of clinical measurement». In: *The lancet* 327.8476 (1986), pp. 307–310 (cit. on p. 48).