Estimation of Ground Reaction Force Based on Computer Vision and Mobile Sensing for Floor Vibration Assessment

by

Yuchen Qian

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Structural Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Floor vibrations caused by human activities, such as walking and running, should typically be addressed during the design phase. However, post-construction evaluations often fail to revisit these vibrations. This gap suggests a need for ongoing assessments to ensure that living experiences align with initial design intentions. To analyze the vibration levels of structures, accurately determining the Ground Reaction Force (GRF) is crucial, as it is one of the most critical components for identifying and predicting floor vibration serviceability. Estimating the floor response to vibrations requires precise input of the excitation load, which involves real-time data from human walking or running activities. In this study, the real-time GRF is estimated using mobile sensing and computer vision methods. The thesis is divided into three key parts. The first part involves developing a Computer Vision (CV) based method to estimate the human walking GRF on a treadmill. This method, based on OpenPose, a deep learning algorithm for multi-person pose estimation, detects key points of the human skeleton. The time-history of displacement from the middle waist is converted into GRF, and the results are compared with those obtained from professional wearable force measurement sensors (Loadsol). The findings indicate that the CV-based method can accurately measure real-time human walking GRF on a treadmill with a root mean square error (RMSE) of 10% in total.

The second part of the thesis demonstrates the use of a smartphone accelerometer to measure GRF. Smartphone accelerometers can record the real-time acceleration of GRF, which can then be converted to real-time GRF. The output GRF is compared with that from the force measurement sensor Loadsol. The results show that the smartphone accelerometer method can accurately record the vertical GRF (with a RMSE of 9.6% in total) when Newton's second law is applied. In the final part, the smartphone mobile sensing method and CV estimation method are evaluated as a cost-

effective and efficient alternative to the Loadsol reference. The thesis concludes with a discussion of the current work's limitations, recommendations, and potential directions for future research.

Preface

This thesis is an original work by Yuchen Qian. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Evaluation of Smartphone and Computer Vision for Walking Load Analysis", No. Pro00141639, Thursday, May 30, 2024. This thesis has been revised with the assistance of Artificial Intelligence models.

Acknowledgements

First, I would like to express my deepest gratitude to my supervisor, Dr. Qipei (Gavin) Mei, and my co-supervisor, Dr. Jianhui Zhou, for their invaluable guidance and suggestions throughout my studies at the University of Alberta. Their support helped me overcome every challenge in my research journey and inspired me to produce a better thesis. I could not have completed this work without their mentorship.

I would also like to thank my supervisory committee, Dr. Ying Hei Chui and Dr. Shay Abtahi for their time and great feedback. Your constructive comments have inspired me to think and dig deeper in my research.

I am also grateful to the members of our research group and my teammates: Hao Xie, Min Hou, Shima Atashgahi, Weijia Zhang, Yue Li, Mahdis Mazhari, Zahra Hassani, Mohammad Talebi Kalaleh, Hui (Polo) Zuo, and Xiao Han. Their assistance and collaboration were vital during my master's degree, and I truly appreciate their support and contribution.

A heartfelt thanks goes to my beloved family and dear friends, Yuyue Zhou, Jinyue Wu, Xiaonan Wang, Xinyu Xu, Xiaowei Yu and who stood by me during the most difficult and trying times. I am incredibly thankful for their unwavering support and companionship.

This research involved human participants, and all experiments were fully approved by the University of Alberta's ethics review board.

Finally, I would like to give thanks to Tim Hortons!! Thanks so much for selling me affordable and tasty coffee during my whole degree, their coffee lets me feel that I am still alive and have the energy to finish the thesis!!!

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Introduction to Floor Vibration Serviceability Assessment

Vibrations resulting from human activities, such as walking, running, and rhythmic movements, can cause significant floor vibrations. Hanagan (1997) confirmed that human-induced Ground Reaction Force (GRF) is a major source of floor vibrations. GRF refers to the human-induced force applied to the floor, which may cause floor vibration and make occupants uncomfortable. These forces occur frequently and cannot be easily isolated. encompasses loads from activities such as walking, running, and cadence movements. Setareh (2010) also claimed that residents are very sensitive to unexpected vibrations; even a small level of vibration caused by walking can annoy the occupants. With the development of efficient materials, slender spans, and lighter-weight constructions, occupants begin to be concerned about excessively human-induced vibrations, even under normal service conditions (Z. O. Muhammad & Reynolds, 2019). While these vibrations do not usually cause structural damage, they can make occupants feel uncomfortable. Therefore, it is necessary to estimate structural vibrations during a building's service life.

There have been multiple complaints about floor vibrations in recent years. For instance, employees at Canada Revenue Agency reported about strong floor vibrations at 441 University Ave. After inspections by three engineering companies, it was confirmed that there was no risk of spontaneous collapse. However, around 300 employees in the Windsor tax building were sent home (CBC News 2019). Similarly, Twenty-One Pilots fans bouncing to the beat at Ottawa’s Place arena caused vibrations severe enough to dislodge ceiling tiles, resembling the effects of damage consistent with a low-level earthquake, (CBC News 2016). There have also been complaints from New York City residents; a family purchased an upscale house with an open floor plan but were bothered by floor vibrations in certain rooms. Other documented cases of unpleasant floor vibrations reported from around the world include iconic structures such as Taipei 101 skyscraper and London's Millennium Bridge (Ashburn and Tech, 2024). Since 1931, when Reiher and Meister began investigating the serviceability limitations of structures, there has been significant progress in understanding and defining human comfort levels related to structural vibrations. Reiher and Meister initially categorized vibrations as 'strongly perceptible' and 'easily perceptible' to describe human sensitivity to structural vibrations. This work marked a milestone in the research on structural serviceability. Those investigations and complaints indicated that floor vibrations can seriously affect residents' lives, with some even unable to work or remain on the premises due to the discomfort.

There remains a gap between controlling vibrations during the design phase and addressing vibration effects post-construction floor because of the lack of adequate design guidelines and fundamental data (Shahabpoor et al., 2017). In other words, the designed vibration and the actual response of a structure may differ due to variations in construction materials and unforeseen human activities. This is largely because the simplified single-human-induced load model is commonly used in dynamic simulation analysis, where the mass ratio of the single-human-induced load is typically too small for accurately estimating vibration comfort. However, most issues with vibration comfort are often caused by crowd-induced excitation (Xie & Hua, 2024). Furthermore, the dynamic parameters of human activities are influenced by various factors, such as weight, height, and the random distribution of crowd loads. The structural response is also affected by the building's function, spatial design, and unpredictable human activity (Helbing et al., 2005).

Previous design codes estimated vertical walking loads without accounting for floor vibration interaction (Racic et al., 2009) and simplified the human walking load into a periodic Fourier series equation, neglecting individual variations in human walking loads and movement (Bucknall et al., 2011). The previous observations are based on factors like floor natural frequencies, the posture of residents, and the mass ratio between occupants and structures in controlled laboratory environment Thus there is a need for real-world human-floor interaction analysis to prevent overestimating floor dynamic responses (Salyards & Noss, 2014). Recent trends in the construction industry have led to the increased the use of large and slender floors in office buildings, which has resulted in damping, now a critical issue in modern structural design. This shift has made dynamic football excitation a governing limit in today's structure design. As a result, discomforting vibrations can occur in the post-construction phase (Z. Muhammad et al., 2019). Therefore, current research should focus on human-floor interaction vibration analysis to accurately estimate the post-construction phase floor vibration. However, before that, GRF must be considered as the most important input value for real-time floor vibration estimation based on individual human walking loads. GRF is fundamental to floor vibration estimation, and the methods of GRF measurement will be introduced in Section 1.2.

1.2 Introduction to Ground Reaction Force Measurement Methods

1.2.1 Direct GRF Estimation methods

There are two main methods to estimate the GRF: direct measurement using force plates and MEMS sensors, and indirect measurement using inertial, acceleration, CV, and motion sensors. For direct GRF estimation, Bocca et al. (2011), and Navabian et al. (2022) confirmed that Micro-Electro-Mechanical Systems (MEMS) sensors are widely used for real-time structural vibration estimation. The MEMS based estimation method requires on-site data collection, but structural vibrations need to be assessed both before and after construction. Therefore, it is essential to evaluate structural vibrations caused by human activities during both phases. Before construction, vibration levels can be estimated using response-based floor vibration design guides, with human activity loads serving as input for analysis. The resulting simulations provide a prediction of the floor's response to vibrations under GRF.

1.2.2 Indirect GRF Estimation methods

For indirect CV-based GRF estimation, marker-based detection is the traditional way to measure human motion, which can then be applied to GRF measurements. Carroll et al. (2013) computed the GRF by tracking 31 markers on objects and compared the results with treadmill-collected GRF, confirming the accuracy of marker-based CV GRF detection. Racic et al. (2010) applied code markers on two objects at 200 Hz to measure the jumping GRF, validating the accuracy of the marker-based GRF detection method by comparing it with force plate data. These methods primarily rely on numerous markers placed on objects, such as optical markers. Nowadays, various marker-free methods are available that can track human motion, such as AlphaPose, BlazePose, and OpenPose. Mundt et al. (2023) noted that AlphaPose and OpenPose have demonstrated reliable accuracy in detecting key points across varying movements, whereas BlazePose is less effective for key point estimation.

Smartphone-based GRF estimations have gained popularity because smartphones are widely used, and their internal accelerometers and gyroscopes are free for most applications on the digital market. This has led to a trend of using smartphones for estimating structural vibrations. Compared to professional sensors, smartphones offer additional benefits due to their small and compact size. They can be easily attached to the human body using tape, allowing them to collect real-time walking or running accelerations and angular velocity in all directions, which can then be computed into GRF for further analysis. This field began with the Nokia N95 mobile device test, equipped with triaxial and uniaxial accelerometers (Lau & David, 2010). Feng et al. (2015) found that smartphones can track the vibrations of a prestressed reinforced concrete bridge with an extremely small error (only 1%) when compared to professional sensors. Ozer and Feng. (2017) claimed that direction-sensitive acceleration data could be corrected using coordinate system transformation, further enhancing the accuracy of the estimation process.

Feldbusch et al. (2017) developed an application called iDynamics, which can record real-time acceleration and angular velocity using the accelerometer and gyroscope inside a smartphone. The accuracy of mobile devices has been validated by comparing them to professional equipment for semi-professional vibration measurement on a bridge. Smartphone vibration estimation has also been used to evaluate dynamic parameters for inspecting concrete bridges, with results compared to finite element models (FEM) of the bridges, demonstrating the potential of smartphones in estimating structural vibrations (Pravia & Braido, 2015).

In recent years, analyzing the reliability of smartphones attached to the human body to collect related GRF has gained attention. To investigate the walking-induced vibration estimation mentioned above, a slender prefabricated prestressed concrete floor was measured using an accelerometer to collect the acceleration data from the lower back as the subject walked across the floor. Martinelli et al. (2020) justified the advantage of smartphones for structural vibration testing, making such surveys cheap and efficient, though still limited in comparison to the accuracy and precision of professional sensors. Smartphones have also been used for vibration serviceability analysis (Cao & Chen, 2020). Smartphone applications and questionnaires were used by volunteers to assess real environmental vibrations and their subjective feelings about the vibration levels, yielding favorable results for smartphone-based estimation (Cao & Chen, 2020). This further demonstrates the reliability of smartphones to monitor vibration levels effectively.

Overall, two new methods—smartphone-based and CV-based—can serve as cost-effective and highly accurate solutions for GRF estimation. These methods allow for real-time GRF data to be imported into Finite Element Modeling (FEM) for floor response vibration estimation.

Before improving the structure design by computing the floor vibration response, the real-time GRF must first be recorded first and then imported into a numerical analysis system for floor serviceability assessment. According to previous research, Chang (1973) identified the perception threshold at 0.1 m/s^2. Subsequent standards, including ISO 2631-1 (Eger et al., 2008) and SCI P354 (Smith et al., 2009) established that perceptible vibration acceleration ranges from 0.015 to 0.02 m/s^2, as confirmed in studies conducted between 1997 and 2009. These findings indicated that people are sensitive to structural vibrations in various types of buildings, including tall buildings and residential structures. Thus, assessing floor vibration performance is essential for enhancing human comfort.

1.3 Problem Statement

The traditional floor vibration comfort assessment system is based on the single-human-induced load model, which does not account for crowd dynamics, variability, or unexpected human activities that contribute to floor vibrations. As a result, there is a gap between the expected floor vibrations in design and those observed in real-world conditions. Therefore, it is necessary to update the methods for estimating human-induced loads.

Previous research has introduced two new methods to estimate GRF compared to the traditional direct force plate measurement. The first method uses a marker-based CV GRF estimation technique, while the second involves attaching a smartphone to the human waist to collect GRF data. However, the first method relies on multiple markers and depth cameras, making it less cost-effective than the smartphone-based GRF estimation method. Meanwhile, the smartphone-based GRF estimation method has not been directly compared to direct GRF estimation, which means it lacks systematic comparison and evaluation. Therefore, this research aims to develop a more cost-effective, non-depth camera-based GRF estimation method and fully evaluate the smartphone-based method by comparing it with the Loadsol system, a widely used direct GRF estimation method.

Systems like Loadsol and other wearable sensors are not cost-effective for large-scale GRF estimation. Specifically, a basic pair of Loadsol insoles costs 2,000 US dollars, and they are highly sensitive to human motion during testing, thus requiring calibration every 30 minutes or less during testing. Additionally, Loadsol insoles are limited by foot size, restricting their use for larger data collection and only allowing GRF measurements on hard floors. In large group GRF estimation scenarios, the smartphone-based method will not be cost-effective enough because all subjects need to use the same brand and version of smartphones. However, the CV-based GRF estimation can detect and analyze multiple subjects simultaneously in real-time, making it a more scalable solution. In this case, CV-based GRF estimation has a better blue picture than smartphone-based GRF estimation. Therefore, the CV-based GRF estimation method combined with a fully evaluated mobile sensing GRF estimation method holds great potential as a cost-effective and time-saving alternative for future GRF analysis.

1.4 Objective and Scopes

The primary objective of this thesis is to comprehensively develop a new marker-free CV-based method and evaluate a cost-effective method (smartphone) for estimating GRF, and to assess their implications for the serviceability of floor vibrations. Specifically, this study aims to:

1. Develop and validate the marker free CV-based GRF estimation method, using the OpenPose library, to process GoPro video recordings, alongside professional Loadsol GRF estimates.
2. Evaluate the performance of smartphone-based GRF estimation methods with those of the professional Loadsol (internal sensor) GRF estimation methods.
3. Analyze and contrast the data collected from smartphone and CV-based estimation methods to identify discrepancies and recommend areas for future improvement.

1.5 Organization of the Thesis

The organization of the thesis is as follows:

Chapter 2 presents the literature review on structural vibration serviceability assessment and GRF estimation methods. It covers the progress of direct and indirect methods for GRF measurement in recent years and the latest achievements in this field.

The chapter 3 focuses on the methodology and experimental design for using CV and smartphones to measure GRF, specifically, implementing and evaluating the effectiveness of the CV algorithm and mobile sensing for GRF estimation.

The chapter 4 focuses on demonstrating how the CV-based method and smartphone-based method can offer scalable and less intrusive alternatives for measuring GRF.

Finally, conclusions, recommendations, and possible future directions are provided in Chapter 5.

CHAPTER 2: LITERATURE REVIEW

2.1 Literature Review on Structure Vibrations Serviceability

The serviceability of structures gained significant attention following the infamous swaying of the London Millennium Footbridge, caused by pedestrian movement in 2000. Located across the Thames River in central London, this incident prompted professionals to analyze the problem in greater detail (Dallard, 2001). Kumar and Kumar. (2014) claimed that various types of structures can be affected by human loading, including floors, footbridges, stadiums, and other buildings. Human activities such as walking, running, jumping, and bouncing can induce vibrations. When these vibrations exceed certain limits, the structures face serviceability issues.

Cao and Chen. (2020) indicated several causes of vibrations: human activity (34.5%), traffic (26.4%), wind (19.1%), machines (12.7%), construction (5.8%), and other sources (1.5%). Their research shows that most structural vibrations are caused by human activity. While these vibrations usually do not cause serious structural damage or failure, they can be problematic if people jump at the same frequency as the structure, potentially causing resonance and structural failure. However, even without causing damage, these vibrations can be felt by people in the building, leading to discomfort or feelings of unsafety. Moreover, people have different definitions of comfort levels. Cao and Chen. (2020) found that individuals with a body mass index (BMI) ranging from 20 to 22 are most sensitive to vibrations. Jones and Saunders. (1972) claimed that women are more sensitive than men at both low and high levels of vibrations. Since a BMI of 20 to 22 is common among adults, structural vibrations can be significant for most residents in a building. Therefore, it is necessary to monitor structural vibrations caused by human activity.

In this century, smart technology is widely used to estimate structural vibrations, such as wireless sensors for collecting vibration data. Estimating structural vibrations is important for assessing structural serviceability, which can optimize a structure's service life. Traditional inertial sensors need to be installed in the structure and transfer data using wires, which involves significant work if the sensors need to be removed, or batteries changed. This process is expensive and invasive. Kim et al. started analyzing these systems in the 1990s. The development of Wireless Sensor Networks (WSNs) provided a low-cost system for structural estimation by integrating sensors and upgrading transmission capabilities (Kim et al., 2007).

Micro-Electro-Mechanical Systems (MEMS)-based sensors reduce installation costs and require less invasive procedures, allowing wireless sensors to be easily installed (Fortino et al., 2012). As a result, wireless sensors have become widely used in the field of structural serviceability estimation. Wireless sensors have four principal components: a sensing interface, signal conditioning section, computational core, and radio transceiver. The transmitter board transfers signals, while the receiver board saves them in computers (Lynch, 2006).

Numerous researchers have demonstrated the reliability of wireless sensors. For example, Jerome Peter Lynch et al. confirmed that wireless sensors offer excellent performance in low-cost solutions for short-term structural vibration acceleration and displacement estimation. However, wireless sensors still have limitations for structural estimation, such as weak high-speed data sampling, high-duty cycles, lack of cost-effectiveness, inability to be stored on a notched surface, and non-wearability (Ruiz-Sandoval et al., 2003). In 2020, wireless sensors were used to collect ground reaction acceleration, which was then converted into GRF (Martinelli et al., 2020). MEMS have achieved wireless data transmission without contact, extending the capabilities of wireless sensors for structural vibration estimation (Sabato et al., 2017). Bridge vibrations were measured by using 113 wireless sensors at multiple points on the South Korean cable-stayed Jindo Bridge, connecting the sensors to a computer to process data and send commands. They found that wireless sensors performed well in estimating bridge vibrations by analyzing cable tension forces (Ju et al., 2015). Whelan et al. (2009) demonstrated the accurate acquisition of vibration acceleration from highway bridges in a real service environment, using wireless sensors set at a sampling rate of 128 Hz collected from 40 channels, achieving high-rate and lossless data collection. Their results confirmed the maturity of wireless sensors to a degree comparable to cable-based systems.

Balageas et al. (2010) stated that MEMS-based wireless sensors perform as well as their macro-scale counterparts but with lower installation costs and less invasive effects. Bocca et al. (2011) demonstrated the high accuracy of wireless sensor structural estimation by comparing time synchronization and high-frequency sample collection results with high-quality wired sensors. Thus, wireless sensors show high performance in estimation, low cost, and easy installation.

Ruiz-Sandoval et al. (2003) highlighted that wireless sensors can address critical structural serviceability requirements, such as synchronized data transmission, dealing with data loss, and identifying damages even with limited collection resources. Navabian et al. (2022) verified that smart wireless sensors can provide accurate time synchronization and high-fidelity structural response with high resolution by testing them on a bridge model excited by a shake table.

Over the past decade, wireless sensors have proven highly effective due to their accurate and high-frequency acceleration recording capabilities. However, practical challenges such as complex installation and removal processes and significant costs hinder their application in structural vibration estimation. These limitations underscore the need for more adaptable and cost-efficient alternatives, prompting the recent shift toward mobile device estimation technologies.

2.1.1 Literature Review on Techniques for Contact GRF Estimation

In recent years, the interest in the evaluation of structural comfort analysis has increased. Wang et al. (2022) noted that although equipment for estimating structural comfort, such as wireless sensors and laser detectors, is expensive, it is simple to install and remove. However, the data collected is typically processed offline resulting in low real-time performance and making it challenging to integrate with big data on the internet. They conducted experiments using an Android device with a self-frequency accuracy of 100 Hz, comparing its estimation results with those from professional instruments, namely a 941B servo sensor and a piezoelectric sensor. The vibrations for the tests were generated by a shaking table that produced three types of vibrations: simple harmonic, white noise excitation, and seismic wave. The data was analyzed using Root Mean Squared Error (RMSE), and the results showed that smartphones could address many limitations of professional instruments, achieving RMSE results of less than 0.2 when compared with the two professional sensors. This suggests that smartphones are particularly effective in estimating structural vibrations, especially when the excited acceleration exceeds 0.05m/s^2.

Previously, researchers fixed smartphones in place to monitor vibration acceleration in a fixed position. Infect, several researchers have verified the reliability of smartphone acceleration estimation on a stable plane. This field began with the Nokia N95 mobile device test, using triaxial and uniaxial accelerometers (Lau & David, 2010). Feng et al. (2015) found that smartphones can track the vibrations of a prestressed reinforced concrete bridge with an extremely small error (only 1%) when compared to professional wireless sensors. Ozer and Feng. (2017) claimed that direction-sensitive acceleration data can be corrected using coordinate system transformation, improving the accuracy of the estimation process. Smartphones have been used in the field of ground vibration acquisition to measure and analyze ground vibrations and determine their effects on structures and subjects, provided the vibration amplitude is above the noise amplitude threshold (Shiferaw, 2021). Smartphones have also proven capable of estimating vibrations in steel bridges and can be used to assess natural frequency under the impact of jumping excitation (Yoon et al., 2013). Sony et al. (2019) summarized that mobile devices can be applied without additional training, are more affordable for big data collection, and can be easily controlled by software platforms remotely to monitor real-time data.

Feldbusch et al. (2017) developed an application called iDynamics, which can record real-time acceleration and angular velocity using the accelerometer and gyroscope inside a smartphone. The accuracy of mobile devices has been validated by comparing them to professional equipment for semi-professional vibration measurement on a bridge. Smartphone vibration estimation has also been used to evaluate dynamic parameters for inspecting concrete bridges, with results compared to finite element models (FEM) of the bridges, demonstrating the potential of smartphones in estimating structural vibrations (Pravia and Braido, 2015).

Zhang et al. (2020) concluded that smartphones can monitor structural behaviour under a shaking table excitation due to the accurate identification of tested three-story bench-scale modal parameters. The accuracy of smartphones was also confirmed by Feldbusch et al. (2017), who verified that smartphone accelerometers can capture structural vibrations well when compared with professional sensors.

Cao and Chen. (2020) confirmed that most floor vibrations are caused by human activity loads (GRF), which are much higher than other natural or transportation-related exceptions. Residents in buildings are particularly sensitive to these vibrations, especially when they are still. Although these loads usually cannot cause structural failure, only the GRF leads to sympathetic vibrations, which may result in floor failure. After construction, this method can also be used to evaluate floor vibrations to enhance occupant comfort and structural serviceability, especially if there are many occupants in the building or if their synchronized jumping movements cause resonance (Tuan & Saul, 1985).

Additionally, smartphones have been used in the field of ground vibration acquisition to measure and analyze ground vibrations and determine their effects on structures and subjects, provided the vibration amplitude is above the noise amplitude (Shiferaw, 2021). Smartphones have also proven capable of estimating vibrations in steel bridges and can be used to assess natural frequency under the impact of jumping excitation (Yoon et al., 2013). Sony et al. summarized that mobile devices can be applied without additional training, are more affordable for big data collection, and can be easily controlled by software platforms remotely to monitor real-time data (Sony et al., 2019).

Kang et al. (2023) also confirmed that smartphones can monitor the acceleration and displacement of two-story structures with a high accuracy of 90%. Feldbusch et al. (2017) developed an application called iDynamics, which estimates structural vibrations using the internal accelerometer of smartphones. They verified the achievements and limitations of this MEMS-accelerometer-based application. They confirmed that the main differences between professional sensors and smartphones are resolution and measuring range. The resolution for professional sensors is up to 10^{-9} g and the measuring range is below negative or positive 1 g, but for smartphones, resolution is commonly from 0.1 to 15 mg with a measuring range of ±4 g. They claimed that higher smartphone sensor resolution reduces noise, improving accuracy. A 100 Hz sampling rate enhances recording quality for all frequencies. Low-resolution devices (10 mm/s^2) have higher noise than high-resolution ones (1 mm/s^2). Amplitude determination requires levels that are well above the noise threshold. Reliable measurements for high-frequency vibrations need sampling rates over 50 Hz. Additionally, Feldbusch et al. (2017) verified that smartphones cannot measure ambient vibrations well because these vibrations are too small. However, they can measure induced vibration effectively, yielding good results compared to professional sensors.

Feng et al. (2015) tested two types of smartphones equipped with MEMS accelerometers. First, they tested them on a shaking table, excited by sinusoidal motions at different frequencies ranging from 0.5 to 20 Hz and discovered that the collected acceleration signal was not perfectly synchronized with the reference signals (PCB Piezotronics sensors), resulting in slight phase differences. They calculated a maximum error of 0.96% in the smartphone frequency acceleration domain estimation. They also tested the capability of smartphones on a large-scale seismic shaking table and a real pedestrian bridge. The results of operational vibration, white-noise excitation, and earthquake excitation confirmed the excellent performance of smartphone vibration estimation, with errors between smartphones and sensors being less than 1%.

In recent years, analyzing the reliability of smartphones attached to the human body to collect related GRF has gained attention. To investigate the walking-induced vibration estimation mentioned above, a slender prefabricated prestressed concrete floor was measured by an accelerometer to collect the acceleration from the lower back when the subject was walking across the floor. Martinelli et al. (2020) first justified the advantage of smartphones for civil structure vibration testing, making surveys cheap and efficient, though still limited by the accuracy and precision of professional sensors. If the vibration level exceeds the human comfort threshold, the structure can be improved by increasing floor slenderness and reducing floor damping (Pavic and Reynolds, 2002).

Overall, previous tests have demonstrated the accuracy of smartphone capabilities in estimating structural vibrations. However, all these experiments were conducted on flat surfaces. A new trend in smartphone vibration estimation involves having humans carry the phone or wearable sensors during the test. This trend focuses on estimating pedestrian-induced GRF. In previous estimations, the GRF can be computed using eighth-order polynomial equations (Sedlacek et al., 2004):

$$\frac{F(t)}{G} = K_1 \times t + K_2 \times t^2 + K_3 \times t^3 + K_4 \times t^4 + K_5 \times t^5 + K_6 \times t^6 + K_7 \times t^7 + K_8 \times t^8 \tag{1}$$

The coefficient K is related to the walking frequency according to the linear regression equations. K varies based on the walking frequency and linear regression equations, which are shown in the following equations from 2 to 9. From k_1 to k_8, they can be computed by linear regression equations, which stands for the walking frequency range from less than 1.75 Hz (Sedlacek et al., 2004).

$$K_1 = -8 \times f_s + 38 \tag{2}$$

$$K_2 = 379 \times f_s - 844 \tag{3}$$

$$K_3 = -2804 \times f_s + 6025 \tag{4}$$

$$K_4 = 6308 \times f_s - 16573 \tag{5}$$

$$K_5 = 1732 \times f_s + 13619 \tag{6}$$

$$K_6 = -24648 \times f_s + 16045 \tag{7}$$

$$K_7 = 31836 \times f_s - 33614 \tag{8}$$

$$K_8 = -12948 \times f_s + 15532 \tag{9}$$

Pedestrian loading is usually computed using the Fourier series with a limited number of harmonic orders (typically less than 5). The Fourier series equation is shown below ((Martinelli et al., 2020):

$$F_p(t) = G[1 + \sum_{n=1}^{N} a_n \sin(n2\pi f_p t + \varphi_n)] \tag{10}$$

In equation 10, $F_p(t)$ stands for the interaction force in real-time series, N indicates the order number of harmonics which is typically less than 5, the G presents the person weight, f_p shows the pacing frequency in Hz, φ_n is the phase angle for the corresponding n-*th* harmonic, a_n is the Fourier amplitude for n-*th* harmonic normalized to person weight.

However, those linear regression empirical formulas cannot accurately reflect the real walking or running load due to various human body walking parameters. Thus, researchers have begun to focus on recording real-time GRF using technical tools such as smartphones, IMUs, and wearable sensors. Martinelli et al. (2020) demonstrated that smartphones with MEMS and triaxial gyroscopes can measure the heel drop test by comparing it with sensors fixed on the floor. They attached the smartphone to the lower back of the subject and compared the acceleration measurements from the smartphone with those from a reference accelerometer on the floor. As a result, smartphones performed well as references. They also tested smartphone GRF estimation by comparing it with reference accelerometers. However, the four reference accelerometers were affected by distance and location on the floor, resulting in about a 30% difference at the farthest reference accelerometer. This discrepancy indicates that smartphones tied to the lower back may not measure real-time GRF as accurately as reference accelerometers.

Therefore, reference sensors are needed to measure the GRF without displacement limitations. Examples include force plates and wearable sensors, which allow humans to walk in the same location or move while walking. McDonald and Zivanovic. (2013) introduced the Constant

Coefficient Method (CCM), a theory that confirms the individual measurement of the total vertical jumping and GRF by the 7th cervical vertebra (C7). This method follows the assumption that the Center of Mass (COM) represents the whole-body mass. By multiplying the acceleration on C7, the GRF can be computed in the time domain. The equation is shown below:

$$GRF_v(t) = m_{total} \times (g + \ddot{x_{v,C7}}(t)) \tag{11}$$

In equation 11, m_{total} indicates the person's weight, g is the gravity acceleration, and $\ddot{x_{v,C7}}(t)$ is the acceleration at the C7 point in real-time series. Based on McDonald's theory and empirical formula, Shahabpoor and Pavic. (2018) applied Opal Internal Measurement Units (IMUs) to measure the GRF, comparing the estimated GRF with a force plate (reference). They found that IMUs collected GRF from C7 and the head with the highest accuracy (0.95 correlation) compared to other body parts. They proposed a Scaled Acceleration (SA) model to process the overestimation of GRF from the CCM theory, thus evaluating McDonald's empirical formula by adding the scale model coefficient. This coefficient can be easily affected by the subject's height, pacing frequency, and weight, and they computed a new formula to calculate γ:

$$\gamma(t) = (GRF_v(t) - m_{total} \times g)/(m_{total} \times \ddot{x_{c,C7}}(t)) \tag{4}$$

They also evaluate the McDonald's equation by adding γ:

$$GRF_v(t) = m_{total} \times (g + \ddot{x_{v,C7}}(t) \times \gamma(t)) \tag{5}$$

In their results, they clarified that the SA method improved the accuracy of GRF estimation by 25% compared to the previous CCM model. They also noted that C7 can measure the GRF accurately after applying the SA model.

In the past, researchers have widely used visual markers to track human movement by dividing key points on the human body, such as on shanks, thighs, hands, and other sections. Racic et al. (2010) used optical markers attached to the skin's surface to build a three-dimensional (3D) representation of movement. They divided the body into 15 parts and located the positions of each mass center to determine the ground reaction force (Bobbert et al., 1991). This research is primarily based on markers on the body. Additionally, we need to compute the GRF from multiple markers (usually 15 to 31), which requires numerous depth cameras and complex computer vision codes to process.

$$F_{GR} = \sum_{i=1}^{S} m_i(a_i - g) \tag{3.1}$$

where m_i is the mass of i th segment of the human body, a_i is the vertical acceleration of i th human body segment, and the g stands for the gravity acceleration (-9.81 m/s^2). The total number of human body segments is 7 in this equation.

Wearable sensors, such as inertial measurement units (IMUs), can be worn on various body parts to collect partial acceleration from each part. Shahabpoor and Pavic. (2018) applied six opal IMUs from C7 to the fourth metatarsal to measure tri-axial accelerations and confirmed that IMUs are powerful tools for measuring GRF by collecting acceleration from C7. They found only a 4-8% root mean square error compared to force plate recordings (GRF). The opal inertial measurement units (IMUs) are also used to measure tri-axial accelerations and orientations at the sternum, fourth metatarsals, lumbar vertebrae, and waist-front. The ground reaction force is computed using Equation 3.2.

$$GRF_v(t) = \sum_{i=1}^{n} m_i \times x_{v,t}\,(\ddot{t}) + g \tag{3.2}$$

Where $m_i$is the mass of segment i th of the human body, $x_{v,t}(\ddot{t})$ is the vertical acceleration, and g is the gravity acceleration. The total number of segments is 13 in this equation. Additionally, humanoid robots can make eye contact and communicate with subjects, recording the head position as the subject walks. Nakanishi and their team built a three-dimensional (3D) head-position model to detect the first-person view of the head pose, allowing the displacement of the human head pose to be tracked during walking or running (Tamaru et al., 2022).

Furthermore, Azure Kinect devices are used for body tracking information collection, which can detect and predict human movement (Posner et al., 2023). The methods mentioned above can track human motion using various equipment, such as IMUs, force plates, and Loadsol. However, this equipment is not cost-effective and is often limited to laboratory environments due to the sensitivity of precise instrument requirements. Therefore, the first experiment of this research uses a more cost-effective and non-laboratory-limited method: CV-based human key-point tracking.

Due to their cost-effectiveness and relative accuracy compared to traditional equipment like force plates, Loadsol sensors have gained widespread use in GRF measurement recently. Insoles can record GRFs for consecutive steps without imposing constraints on foot placement. By comparing the validity and repeatability of three types of insoles (Pedar, Medilogic, and Tekscan), Price et al. confirmed that the Pedar insole has the best accuracy and repeatability (Price et al., 2016).

Insole-based estimation of the vertical GRF was compared to force plates using Gaussian process regression from two walking steps, estimating vertical ground reaction force with mean errors of 8% (Eguchi et al., 2020). For walking, 11 sensors are considered suitable, and for jogging, between 7 and 11 sensors are suitable (Fuchs et al., 2022). A Long Short-Term Memory (LSTM) model was developed for predicting 3D GRF, and after integration, it can reflect real-time forces (Hajizadeh et al., 2023).

The Loadsol, investigated by Novel Company, is a powerful wearable insole sensor that measures GRF by integrating all the pressure from the insole and converting pressure to force in Newtons. The novel nanocomposite piezo-responsive foam (NCPF) developed by Parker et al. is a wearable sensor that can estimate the 3D GRF during human walking with a mean average error of 2.15% in the vertical direction (Rosquist et al., 2017). Kim et al. (2020) also designed a two-dimensional force sensor (M2D) wearable GRF estimation sensor, demonstrating high accuracy in measuring vertical GRF, with errors ranging from 2.52% to 3.1%.

Seiberl et al. (2018) tested the GRF estimation capabilities of a new wireless insole (Novel Loadsol) force sensor by comparing it with a gold standard reference (force plates). They measured vertical GRF simultaneously with force plates (1 kHz) and Loadsol sensors (100 Hz) during running movements among 10 subjects. The results showed that the main bias for ground contact time, peak load, and impulse ranged from 0.6% to 3.4%, demonstrating the high performance of Loadsol in capturing the dynamic behaviour of GRF.

Weizman et al. (2019) used Pedar insoles and Smart insoles to measure the Center of Pressure (COP), validating the results against a reference device (Kistler force plate). They noted that the RMSE showed a strong correlation between Pedar insoles (5.6%) and the force plate. Additionally, the coefficient of determination for Pedar insoles was 0.9964, indicating a good correlation between Pedar insoles and force plate COP measurements.

Fong et al. (2008) demonstrated the high accuracy of the Novel Pedar insole for multi-directional GRF estimations. They tested the correlation between the Novel Pedar insole (100 Hz) and a force plate (1000 Hz) for anterior-posterior, vertical, and medial-lateral direction GRFs. The correlation coefficients were 0.928, 0.898, and 0.719, respectively, proving that Pedar insoles can independently measure GRF in any direction, even outside a laboratory environment.

Kong et al. (2023) validated that the Pedar Loadsol is useful for in-field measurement. They tested the GRF collected from Pedar Loadsol and an instrumented treadmill on flat, inclined, and declined surfaces, confirming that Pedar Loadsol can measure GRF accurately even when multiple subjects were carrying heavy loads on the treadmill. The Bland-Altman and Limits of Agreement (95%) methods were applied for insole validation, showing strong similarities between the insole and instrumented treadmill tests. Using the intraclass correlation coefficients (ICC) comparison method, the validity of ICC results ranged from moderate to excellent (0.686–0.982). Peebles et al. (2018) claimed that the repeatability ICC results from Loadsol sampled at 100 Hz and 200 Hz were 0.686-0.982 and 0.765-0.987, respectively.

Peebles et al. (2018) introduced that the Novel Loadsol can compute GRFs with high accuracy by comparing it with an instrumented treadmill during walking and running. The intraclass correlation coefficients (ICCs) between Loadsol and the treadmill ranged from 0.88 to 0.96, while the reliable across-session ICCs ranged from 0.00 to 0.03. These results indicate that the novel insole pressure measurement is a valid tool for GRF estimation. Additionally, frequent calibration and a rapid decrease in sensitivity for insole running tests make insoles less reliable (El Kati et al., 2010).

The Loadsol features a unique, flexible sensor that covers the entire plantar surface, capturing forces exerted between the foot and shoe at any contact point. This sensor employs a patented technology for linear measurement, enabling precise quantification of partial loads. By integrating data from 99 sensors embedded in the Loadsol, the pressure generated during human walking can be accurately converted into ground reaction force (GRF).

The accuracy and reliability of Loadsol have been confirmed by previous research. Price et al. demonstrated that Loadsol has excellent validity, with a Root Mean Square Error (RMSE) of 2.6 kPa and a 3.9% difference compared to force plates (Price et al., 2016). Additionally, Kong et al. highlighted that Loadsol's high accuracy aligns well with standard laboratory equipment, such as instrumented treadmills. They affirmed that Loadsol is an effective tool for measuring GRF during walking (Kong et al., 2023) In summary, various technical tools, including sensors, Novel Loadsol, Pedar Loadsol, IMUs, and smartphones, have demonstrated high accuracy in GRF estimation. However, some of these tools are not cost-effective for widespread use, indicating the need for more economical methods in future developments.

2.1.2 Literature Review on Contactless Techniques in GRF Estimation

Nowadays, computer vision has emerged as a promising method for GRF estimation due to its cost-effectiveness and high-frequency estimation capabilities. Computer vision techniques can be broadly categorized into two approaches: the first involves using markers attached to subjects, while the second utilizes marker-free detection for GRF estimation.

For marker-based GRF estimation, Racic et al. (2009) employed a motion-capturing system based on video optoelectronic technology to verify position and displacement with a three-dimensional reproduction of human body movement. They used Coda motion technology to process the GRF estimated by the video optoelectronic system and validated it against GRF collected by a force plate during jumping movements. The results from the two methods showed that the first dominant harmonic of the Fast Fourier Transform (FFT) processed GRF had less than a ±2 percent difference compared to the force plate recordings. The difference for the second harmonic GRF was within 4%, and for the third harmonic GRF, it was less than 5%. These results demonstrate a good match between force plate recordings and the reproduced GRF in the frequency domain.

For marker-free detection tools, Albert et al. (2020) utilized a Microsoft Kinect camera (Azure Kinect) to assess tracking performance by comparing it with a gold standard Vicon multi-camera system and a 39-marker plug-in Gait model. They confirmed the high performance of Azure Kinect for tracking. However, there are several open-source tools available on GitHub that can serve as alternatives to this depth camera for human motion tracking, such as OpenPose, AlphaPose, BlazePose, and OpenCap.

Mundt et al. (2023) reported that OpenPose and AlphaPose are two highly effective computer vision tools for GRF estimation. They applied multiple camera-based systems, including OpenPose, AlphaPose, and BlazePose, to detect key points on a running human body. Their findings revealed that AlphaPose (98.4%) and OpenPose (94.5%) achieved similar accuracy in key point detection. In contrast, BlazePose exhibited lower accuracy (65.2%), missing approximately half of the key points during detection. They utilized Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) methods to process the key points, comparing them with a force plate reference. The resulting RMSE values for GRF were 0.36 for OpenPose and 0.34 for AlphaPose.

Verheul et al. (2024) used a new marker-less motion capture tool, OpenCap, to estimate GRF during human jumping. When validated against a gold-standard force plate, they observed that the bias and limits of agreement between OpenCap and the force plate ranged from 5% to 15%. They also determined that a variable bias of less than 15% can be considered indicative of good performance for detecting changes in GRF variables between different jumping movements.

In summary, previous research has demonstrated that computer vision-based methods can be highly effective for human motion and GRF estimation. However, earlier CV detection methods predominantly relied on markers or neural network training. Consequently, there is a need for tools that do not depend on depth cameras, markers, or neural network learning to estimate GRF in non-laboratory environments.

CHAPTER 3: METHODS AND EXPERIMENT DESIGN

3.1 CV-based GRF estimation

Traditional marker-based methods typically require 15 to 31 markers attached to the subject's body (Bobbert et al., 1991) and using depth cameras for displacement tracking. For multi-marker GRF estimation, the human body mass is divided into several segments, and the GRF of each segment is integrated to calculate the overall GRF (Shahabpoor & Pavic, 2018). However, this approach can introduce errors in measuring the weight of individual body segments and relies on numerous markers, complicating the measurement and processing of GRF. To address these challenges, this thesis develops a marker-free CV-based GRF estimation method that eliminates the need for depth cameras, body segment mass measurements, and markers, thereby simplifying the process and enhancing cost-effectiveness.

The method utilizes the OpenPose library, which is based on Convolutional Neural Networks (CNN). The reliability of the method is evaluated by comparing its results with those obtained from the professional Loadsol reference.

3.1.1 Loadsol Reference for CV-based GRF Estimation

In this research, the Novel Loadsol will be used as the reference for GRF measurements. Loadsol sensors have recently gained popularity for GRF measurement due to their cost-effectiveness and relative accuracy compared to traditional equipment, such as force plates. These insoles can record GRFs for consecutive steps without imposing constraints on foot placement. Extensive data experiments by Price et al. (2016), Seiberl et al. (2018), Fuchs et al. (2022), Hajizadeh et al., (2023), and Kong et al. (2023) have all validated the high performance of Loadsol for GRF estimation. Therefore, the Novel Loadsol will be utilized as the reference for evaluating the CV-based GRF estimation method.

Figure 3.1. Flow The Novel Loadsol Overview

In summary, Loadsol demonstrates high performance in GRF estimation and will be used as the reference in this thesis. The Loadsol system supports sensor scanning at frequencies of up to 200 Hz, providing detailed insights into the force dynamics under the foot. A key feature of this system is its smartphone app, which delivers immediate auditory and visual feedback based on the collected force data. The app also offers comprehensive analytics on parameters such as peak force, cadence, loading rate, contact time, and symmetry.

3.1.2 The Methodology of OpenPose Detection

OpenPose is a real-time multi-person detection system developed by the Perceptual Computing Lab at Carnegie Mellon University. It is capable of tracking multiple human body parts, including the face, hands, neck, and feet, and is widely used in both computer vision and machine learning fields (Cao et al., 2021). OpenPose operates based on real-time multi-person 2D motion estimation, which involves detecting key points of individuals in images and videos. This method employs a nonparametric representation known as Part Affinity Fields (PAFs) to learn and identify body parts within images. Additionally, the algorithm leverages Convolutional Neural Networks (CNNs) to enhance the accuracy of body part observations.

The primary neural network underlying OpenPose is based on Convolutional Neural Networks (CNNs). In CNNs, each unit in a layer is connected to a small neighborhood of units in the previous layer, allowing for the extraction of local features such as edges and corners. For feature maps, sets of units with identical weights consistently detect features across all images, making the system resilient to shifts and distortions. Subsampling layers further enhance this robustness by performing local averaging and reducing resolution, which minimizes sensitivity to minor shifts and distortions. Consequently, CNNs can be efficiently implemented in hardware, achieving high speeds for tasks such as character recognition (LeCun & Bengio, 1995).

The core of this method employs the CNN-based OpenPose library to process video recordings of subjects walking, captured using a standard camera (not depth-based). OpenPose detects key point displacements in two dimensions, particularly tracking waist displacement, which is then converted into real-time acceleration. By multiplying this acceleration by the subject's total body weight, the GRF is calculated. This computed GRF is compared with the results from the professional Loadsol estimation to evaluate the accuracy of the method.

For detection applications, OpenPose is highly effective as it can simultaneously detect multiple key points and distinguish between individuals even in crowded environments. Beyond full-body detection, OpenPose offers specialized models for tracking various key points with detailed precision. As open-source software, OpenPose is accessible for integration into individual projects focused on human movement tracking. Unlike other computer vision methods, OpenPose excels in real-time detection of multiple subjects within the same image or video, accurately identifying and tracking different body parts independently for each subject independently.

OpenPose can detect 25 key points on the body, including hands, ankles, arms, face, and feet. It provides real-time performance for multi-person 2D pose detection, which is essential for accurate human walking estimation. The algorithm is built upon convolutional neural networks (CNNs) and relies on two main components: heatmaps and part affinity fields (PAFs). Heatmaps represent the confidence levels of detecting each body part at every pixel, while PAFs indicate the degree of association between different body parts, enabling accurate differentiation within the same image.

PAFs have been employed in previous models for benchmarks such as the COCO2016 key points challenge and the MPII Multi-Person benchmark due to their accuracy and efficiency. This method uses a two-branch CNN to predict heatmaps and assemble body parts into whole-body poses (Cao et al., 2021). Wei et al. (2016) advanced pose estimation by integrating Convolutional Pose Machines (CPMs), which combine the strengths of convolutional networks and Pose Machines. This approach effectively processes diverse poses and views from multiple angles.

For object manipulation or situations involving multiple interactions, OpenPose can still generate geometrically consistent annotations, addressing challenges in scenarios prone to occlusion (Simon et al., 2017).

3.1.3 The Process of OpenPose GRF Estimation

The flow chart of the CV method is shown in the figure below:

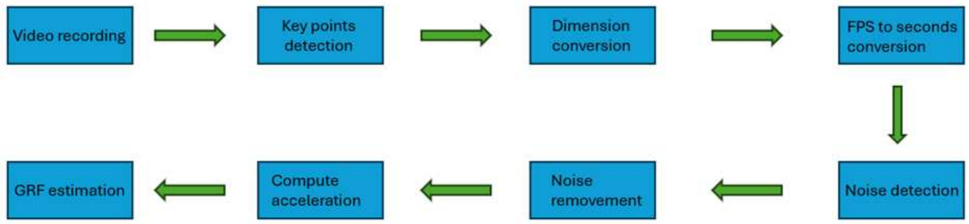

Figure 3.2. Flow chart for OpenPose GRF estimation process.

In In this experiment, the subject's walking motion is recorded using a GoPro Hero 9 camera. Prior to utilizing OpenPose for tracking, a distortion check is performed to ensure overall detection accuracy, addressing common issues related to camera distortion.

- The subject will walk on a treadmill, with the GoPro Hero 9 mounted on a tripod to record the subject’s motion.
- The recorded motion video will be imported into the OpenPose library for key point displacement detection, measured in pixels at this stage.
- After walking on the treadmill, the subject will hold a chessboard, and the real-world dimensions of the chessboard will be recorded for distortion check purposes.
- The GoPro camera's frame rate is set to 240 fps, capturing 240 frames per second, with a resolution of 1080P in this experiment. This information will be used to convert fps to time in seconds.
- Due to the high frequency of the GoPro and potential detection errors at each detection step, there may be high-frequency detection noise that could affect the key point displacement results. Therefore, stationary noise detection will be conducted before converting the displacement to GRF.

- The high-frequency noise can be removed using a filter that only allows low-frequency components to pass through, thereby eliminating the high-frequency noise.
- After noise removal, the real-world displacement will be converted to acceleration using Newton's second law. The GRF can then be calculated by multiplying the acceleration by the body mass.
- Finally, the GRF will be computed.

There are two main types of distortion (OpenCV, n.d.). Retrieved from https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html

- Radial distortion, which causes straight lines to appear curved
- Tangential distortion occurs when the lens is not aligned with the image plane, making objects appear closer than they are.

Both types of distortion will be explained and checked in the experiment section of this chapter.

3.2 Smartphone Based GRF Estimation

To investigate the capabilities of smartphones for estimating GRF, mobile devices are attached to the subjects' waists (the central body mass) to collect acceleration data. This acceleration data is then converted to GRF using Newton's second law. The estimated GRF from the mobile devices is compared with the reference measurements from Loadsol to evaluate the reliability of mobile device accelerometers for GRF estimation.

This thesis will introduce a new approach using mobile device apps to estimate the GRF in various scenarios, including different applications, specific walking frequencies, and individual weight parameters. The results will be compared with Loadsol, a cutting-edge technology that provides real-time GRF measurements.

In this experiment, two Samsung Galaxy S21 smartphones (Model SM-G991W) are used for GRF estimation. These smartphones are equipped with MEMS accelerometers with a sensitivity range from ±2g to ±16g. The applications employed are MATLAB Mobile and Phyphox, which can connect to a computer or laptop (MathWorks, n.d.). The Y-axis (as shown in Figure 3.3) is oriented longitudinally to the screen and perpendicular to the ground. For the experiment, only the acceleration data in the Y-axis (excluding gravitational acceleration) will be utilized for analysis. The smartphones are securely attached to the subject's front waist and ankle using transparent adhesive to prevent any movement or slippage during the test.

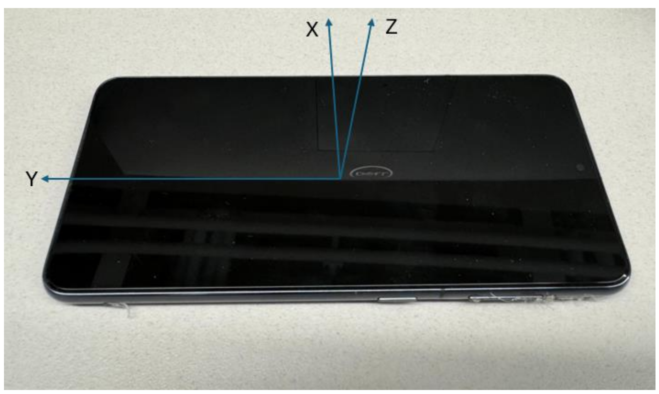

Figure 3.3. Orientation sensors log data in relation to the X, Y and Z axes.

During the test, subjects walk at a comfortable pace to acclimatise themselves to the smartphone. GRF acceleration data is recorded in real-time throughout the walk. The subjects walk on a flat floor in an office building, free from external disturbances, following only the start and stop commands. Walking begins when the MATLAB code is initiated and the accelerometer starts recording acceleration data, synchronized with the start command.

The test is repeated three times for each subject, and the most stable and consistent results are selected for further GRF analysis.

3.3 Experiments

3.3.1 CV

3.3.1.1 Experiments for CV-based Camera Distortion Check

Before recording the real-time walking motion of the subject, it is essential to check camera distortion to ensure the lens is perfectly parallel to the ground. Two types of distortions need to be addressed: radial and tangential distortion. These distortions can be computed using the following equations:

The radial distortion can be calculated using the following equations (OpenCV, n.d.):

$$x_{distorted} = x(1 + k_1r^2 + k_2r^4 + k_3r^6) \tag{3.3}$$

$$y_{distorted} = y(1 + k_1r^2 + k_2r^4 + k_3r^6) \tag{3.4}$$

The tangential distortion can be determined by the equations below (OpenCV, n.d.):

$$x_{distorted} = x + [2p_1xy + p_2(r^2 + 2x^2)] \tag{3.5}$$

$$y_{distorted} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \tag{3.6}$$

Where k_1, k_2, k_3 stand for the distortion coefficients, $p_1$and $p_2$stand for the tangential distortion parameters. For radial distortion, $x_{distorted}$ and $y_{distorted}$ are distorted coordinates, x and y are original coordinates in the picture, r stands for the distance from the center of the image (radial distance). The process of flow chart is shown in figure below:

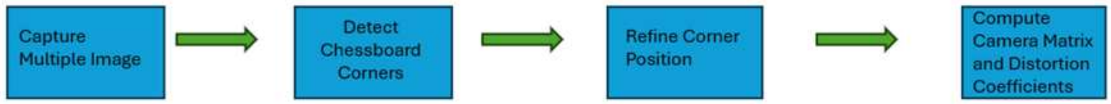

Figure 3.4. Flow chart for checking camera distortion.

To achieving this progress, 20 chess board pictures are used for distortion coefficient calculations, the pictures are shown below:

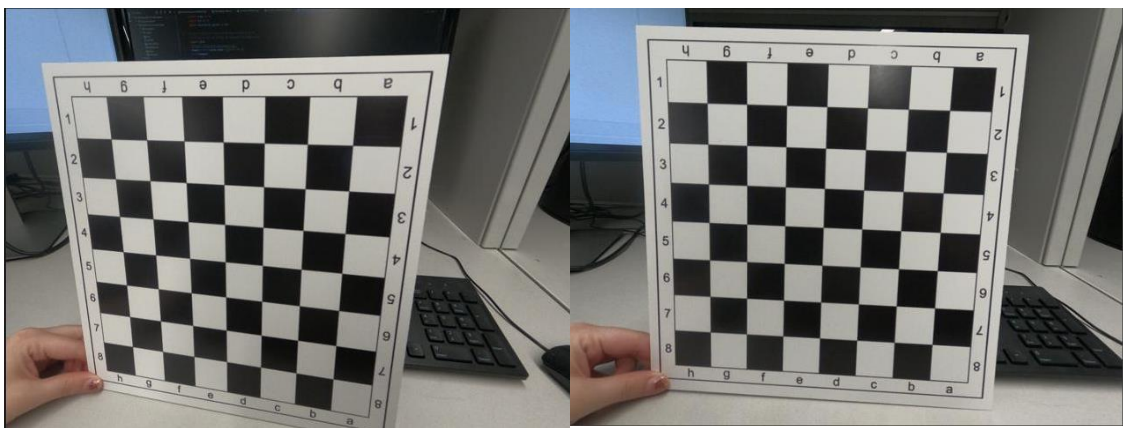

Figure 3.5. Pictures for checking camera distortion.

To perform a camera distortion check, key information including the dimensions and cell size of the chessboard is essential. Specifically, the number of inner corners per row and column of the chessboard (e.g., 8x8 in the figure above) must be known. Additionally, the size of each square on the chessboard is required to ensure accurate distortion check results. These parameters are crucial for detecting the chessboard pattern in images and calculating the camera's intrinsic and extrinsic parameters accurately.

In the distortion check process, relative points on a chessboard are crucial. The corners of each chessboard square are used as reference to establish points for real-world coordinates. A minimum of ten test images is required. The input data for the distortion check consists of a set of 3-dimensional real-world coordinates and the corresponding 2D image coordinates, where the chessboard is kept in the X and Y planes, with Z-axis values consistently set to zero. Figure 3.6 shows the original image and the image with circles overlaid on the corner coordinates.

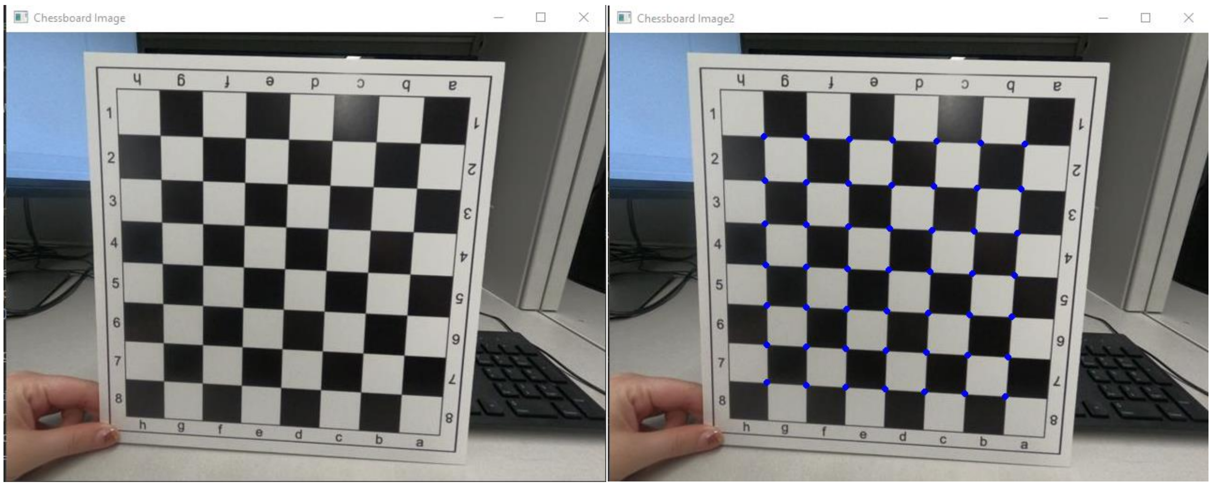

Figure 3.6. Original image (on left) and image (on right) with circles overlaid on corner coordinates.

The image is then processed using a grayscale filter provided by OpenCV, which refines the corners to sub-pixel accuracy. Figure 3.7 shows the comparison between the original corners and the refined corners in the grayscale image.

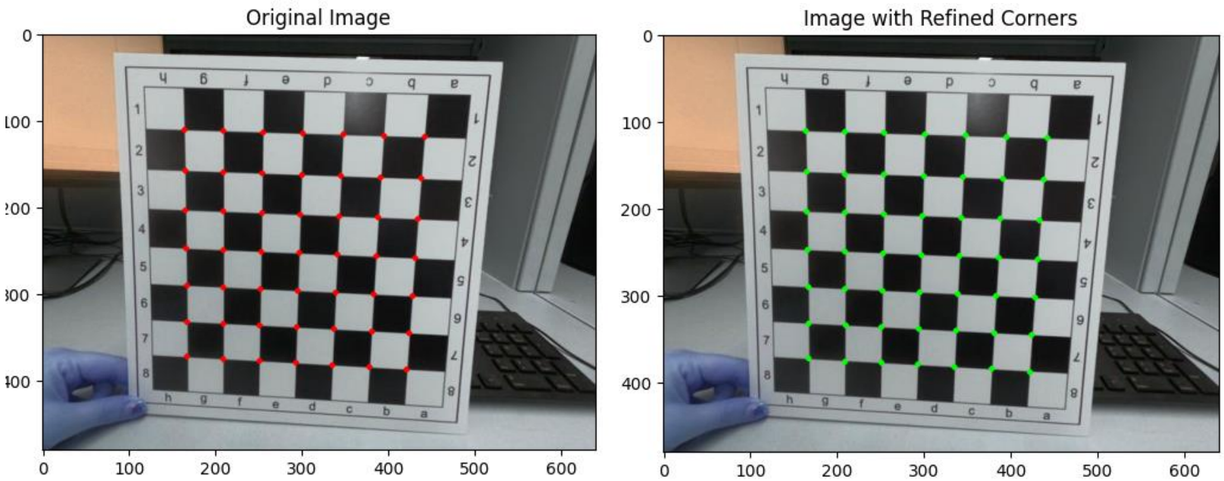

Figure 3.7. Comparison between original corners and refined corners.

After the corner refinement process, the refined corners are displayed as red circles if the board is not found, or as colored corners if detected successfully, as shown in Figure 3.8.

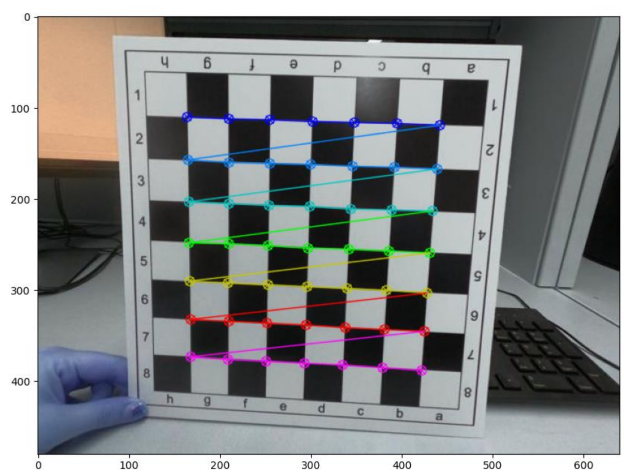

Figure 3.8. Display of corners.

In this distortion check process, 20 images were used for training and testing. The distortion matrix for the GoPro Hero 9 indicated minimal distortion, and the distortion parameters were found to be insignificant. As shown in Figure 3.12, there is no noticeable distortion correction, and the output distortion coefficients are small enough to be considered negligible.

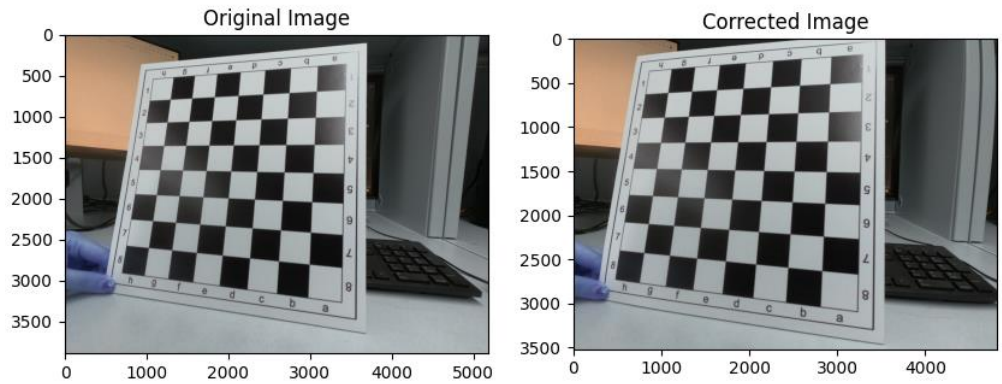

Figure 3.9. Comparison between raw picture and corrected picture by applying calibration.

The camera matrix and distortion coefficient are shown in the following matrix.

$$Camera\ Matrix = \begin{bmatrix} 303.116 & 0 & 318.459 \\ 0 & 304.868 & 240.835 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Distortion\ Coefficient = [0.019 \quad -0.032 \quad 0.001 \quad 0.000 \quad 0.025]$$

To minimize potential distortion, it is important to keep the subject centered within the video frame, as distortion primarily occurs at the corners. While this distortion check process can also be applied to calibrate video distortion, it is not required in this case since the distortion matrix is specific to each camera. Therefore, recalibration during the video recording process is unnecessary. In summary, the distortion coefficients confirm the accuracy of camera detection, making further distortion calibration redundant.

3.3.1.2 Experiments for CV-based GRF Estimation Setup

Computer vision is applied using the open-source code OpenPose, which tracks human motion by identifying key points on the human body. These key points are highlighted and connected by colored lines, as illustrated in Figure 3.10. The coordinates of each key point can be extracted from this motion detection process. Notably, the coordinates for the neck and waist are not derived from real-time position data but are instead calculated as the average coordinates of the two shoulders and thighs. Despite this, the coordinates for the neck and waist are considered stable and reliable for body coordinate measurements.

Figure 3.10. Key points detection of subject.

At the beginning of the test, the subject walks along a path parallel to the camera, simulating a typical indoor walking scenario. A GoPro camera is positioned in front of the subject to record their walking motion. However, using a non-depth camera like the GoPro presents a challenge. In the images, horizontal features at the bottom and vertical lines do not align as expected with the camera model. Specifically, the horizontal features at the bottom of the images suggest that the camera is tilted upward. This tilt causes a distortion in the photo, affecting the apparent size of the features in the images (Jin et al., 2023).

In this experiment, the subject is asked to walk on a treadmill to ensure that all movements occur within the same dimension and are not affected by changes in lighting or camera depth issues. The camera depth issue can cause significant decreases in amplitude, which can affect the accuracy of acceleration derived from real-time displacement. This issue could be influenced by multiple factors. First, vertical displacement from the subject to the GoPro camera is inconsistent. Although the subject walks in a linear direction, the distance between the camera and the subject is not stable all the time, as the camera remains stationary. This can cause slight perspective distortion due to the lens curvature and angle of view in the camera field. The second factor is that OpenPose can only analyze posture or motion in a 2-dimensional field, which can result in incorrect depth information and affect displacement tracking. These issues cause incorrect displacement measurements in the experiment, making this method unsuitable for further consideration. To accurately track key points, the subject is required to wear shorts or sports tights. The camera is positioned approximately one meter away on the right side of the subject to capture all walking movements at the center of the frame, where distortion is minimized, as distortion typically occurs at the corners of the video frame.

3.3.1.3 Experiments for CV-based GRF Estimation Key Points Units Conversion

The output from tracking the key points of the subject's walking is measured in vertical displacement in units of pixels. However, pixels cannot be directly used for acceleration analysis and must be converted into real-world dimensional displacement. To facilitate this conversion, the subject holds a chessboard with known dimensions before the experiment begins. The conversion factor is then calculated by establishing the relationship between the real dimensions of the chessboard and the corresponding dimensions in pixels, allowing for accurate computation of displacement. The actual dimensions of the chessboard are 276 mm in width and 274 mm in height. The conversion factor can be computed using the following equations (GitHub, 2023):

$$Width\ Conversion\ Factor = real\ width\ \div width\ in\ pixel \tag{3.7}$$

$$Heigh\ Conversion\ Factor = real\ heigh\ \div heigh\ in\ pixel \tag{3.8}$$

These formulas can translate pixel measurements from digital images into actual physical dimensions. This process is essential for object detection. The assumptions are:

- Planar Movement: these two formulas assume that the objective movement is parallel to the GoPro camera, in other words, there is no movement along the depth axis, such as moving toward or away from the GoPro Camera.
- Known Real-word Dimensions: The real-world dimensions which include the width, and the length of chessboard must be known correctly. These real-world dimensions information is the reference for the conversion factors.
- Camera Distortion Check: The camera must be calibrated before conversion to ensure that the pixel dimensions accurately reflect the scene.

The large difference in displacement detection when the object is moving away is caused by movements in a plane parallel to the Complementary Metal-Oxide-Semiconductor (CMOS) sensors. These assumptions can address the issue of decreased acceleration. The large difference in the previous floor walking test is caused by the object's size change during recording, leading to errors in displacement estimation and dimension conversion.

To find the dimensions in pixels, a Region-of-Interest (ROI) is introduced into the conversion process. ROI is a subset of an image that focuses on a certain area for further analysis, such as zooming in on a part of the image that researchers are interested in. In this experiment, ROI is used in the computer vision field for chessboard detection. ROI can detect where the object is and measure its dimensions in pixels.

The chessboard is selected using the OpenCV algorithm, which focuses on the chessboard throughout the entire image. This isolates the chessboard from the rest of the image and potentially improves the accuracy of the analysis. Additionally, ROI performs well in edge detection, contributing to the precise visualization of dimension data.

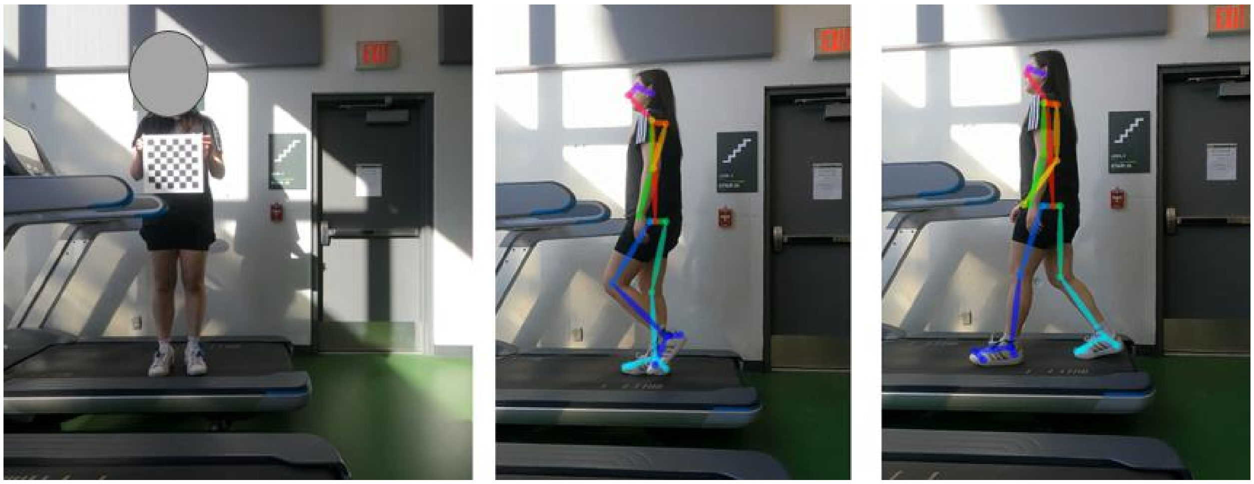

Figure 3.11. OpenPose GRF estimation test on treadmill.

The video frame rate is set to 240 frames per second (FPS) or 240 Hz, with a video resolution of 1080P during the test. The FPS stability is confirmed before the test. For the reference load measurement, Loadsol is used with a sampling frequency of 100 Hz.

Next, the walking motion recording is processed using the OpenPose algorithm to detect key point displacements. The displacement of the waist key point is then converted into real-time acceleration. By subtracting gravitational acceleration and multiplying it by the subject's body weight, the GRF estimated by OpenPose is compared with the Loadsol reference.

3.3.2 Smartphone

3.3.2.1 Smartphone-based Experiments Setup

For this study, the reference data is collected using Loadsol and the Loadsol application. Traditionally, treadmills and force plates have been used for GRF reference, and their accuracy has been validated by multiple experiments. Analysis of gait using a treadmill has been found to be functionally equivalent to evaluating overground gait (Riley et al., 2007). However, while these two types of equipment are reliable for high-frequency estimation and allow individuals to choose their walking speed, they are confined to laboratory environments and are not cost-effective for broader applications.

Figure 3.12. Immobilization of smartphone for GRF acceleration estimation.

For smartphone accelerometer estimation, the primary challenge lies in the coordinate system and rotation correlation. The smartphone cannot remain perfectly vertical while humans are walking due to the cyclical shift in body weight from the left foot to the right foot. This results in the body tilting at various angles throughout the walking cycle, regardless of walking frequency. To address this issue, angular velocity is incorporated into the estimation process. Time-domain angular data is extracted from the gyroscope sensor and combined with the acceleration data in the vertical direction to improve accuracy.

The goal is to isolate the true vertical acceleration from measurements that may be influenced by features in other orientations. In MATLAB, acceleration is typically measured along three perpendicular axes, labelled X, Y, and Z corresponding to the device's orientation-based coordinate system.

3.3.2.2 Smartphone-based Experiments Angular and Time Step Correction

There are two technical issues that need to be addressed. The first is vertical acceleration, which may include components from other directions during data collection. The second is time synchronization, as the two phones might have different time steps, leading to delays or advances during the test. Both issues must be corrected before proceeding with further analysis.

For vertical GRF acceleration correction, only the acceleration in the vertical (Y) direction is needed because the GRF collected from the Loadsol reference monitors the load in the true vertical direction. However, during human walking, weight shifts from the left foot to the right foot, causing changes in the center of mass. As a result, the subject cannot always maintain a perfectly vertical orientation relative to the ground. This leads to some rotation and tilt during walking, causing orientation changes and rotations around one or multiple axes. These rotations can cause the accelerometer to record acceleration that is not aligned with the subject's movement in specific directions along certain axes.

To correct the collected vertical GRF acceleration, the angular velocity should be recorded simultaneously to measure the rotation and tilt during the subjects' walking. The angular velocity indicates the rate of rotation around the corresponding axes. By integrating the angular velocity over time, the rotation of each axis can be computed. Trigonometry and correction functions are then used to determine the relationship between angles and side ratios. The mathematical function used in this correction is shown below:

$$Y_{corrected} = Y \times cos\ (Angle_Y) - X \times sin\ (Angle_Y) - Z \times sin\ (Angle_X) \tag{3.1}$$

The theory for vertical (Y) direction acceleration correction is that the angle of rotation is used to ensure that they correctly represent the movement around the certain axes and adjust for tilt and rotation in X, Y, and Z direction respectively. The recorded vertical acceleration is corrected by using equation 3.1, after that the pure vertical acceleration is saved for further analysis. Equation 3.1, $Y_{corrected}$ presents the corrected value of acceleration in the Y direction after accounting for angular adjustments, X, Y, and Z all stand for the principal directions along the screen. The is the cosine of the angle associated with the Y coordinates, this term adjusts the acceleration in the Y direction based on its angular displacement, and the same for the rest of the angular adjustment process. The raw phone estimated GRF and corrected GRF are shown in figure 3.13.

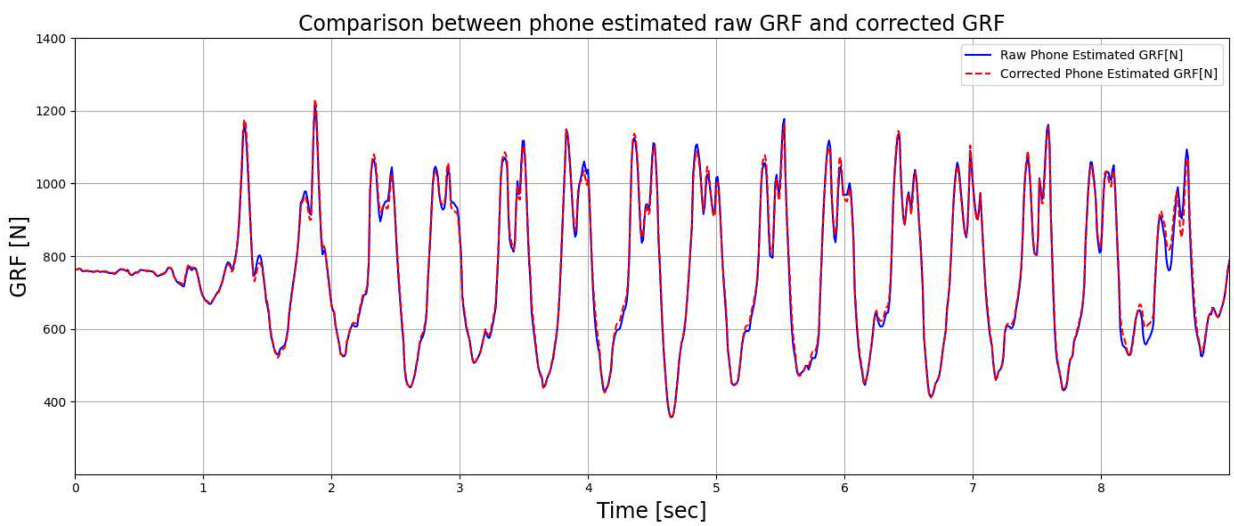

Figure 3.13. Comparison of GRF between before and after angular adjustment.

For time synchronization correction, it is essential to ensure that all smartphones are set to the same time zone. The synchronization of the two phones is achieved via General Packet Radio Service under the same mobile network. The screenshot of the two same-system mobile devices is shown below. From the time zone readings, it is evident that both devices are in the same time zone, controlled by internet time synchronization.

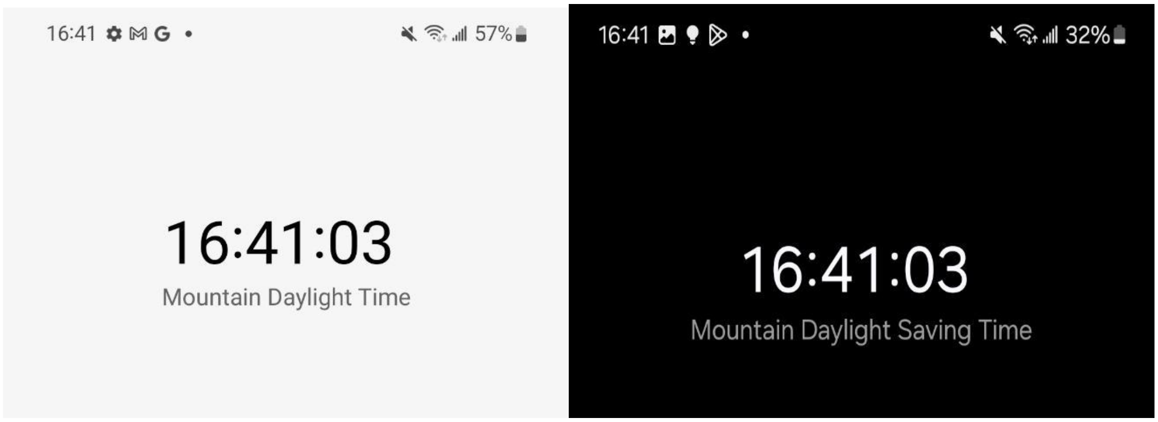

Figure 3.14. The synchronization adjustment screenshot.

For this experiment, the Loadsol is calibrated using the calibration function in the mobile application. The subject is asked to first unload one foot, then reload onto the Loadsol, and finally unload the foot again. After calibrating one foot, the same process is repeated for the other foot. This process is repeated several times for both feet until Loadsol reflects the subject's body weight within a tolerance of 5% difference. Once calibration is complete, the GRF data can be exported from the application. The result of the GRF at a comfortable walking frequency is shown in Figure 3.15.

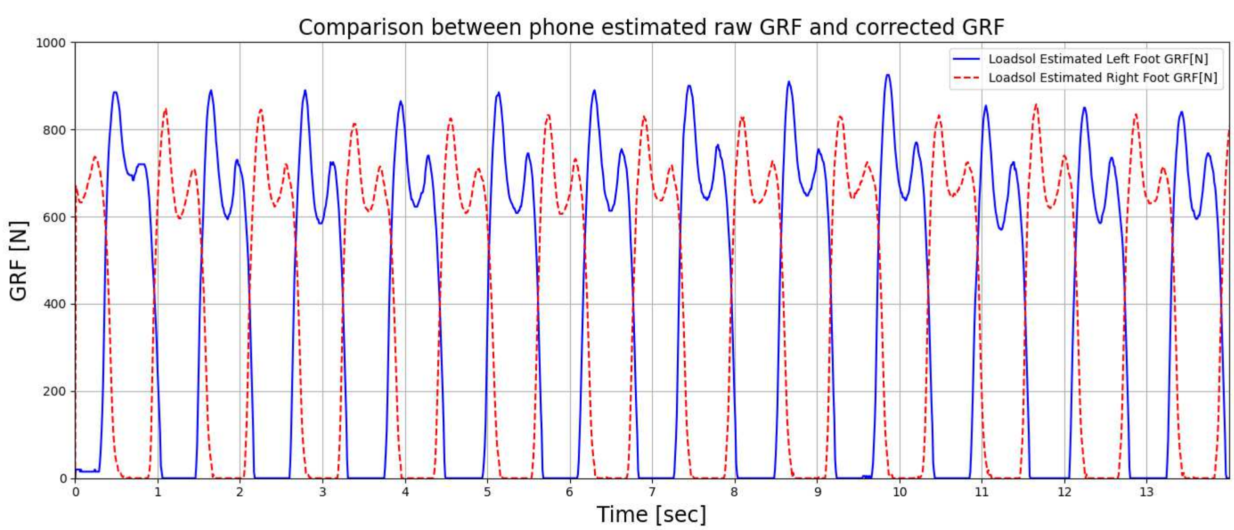

Figure 3.15. Loadsol GRF reference results.

The waist is one of the most stable parts of the human body. Since high-frequency signals from other parts can interfere with estimation, it is necessary to choose a stable body part to minimize these high-frequency signals. In this experiment, the load collected from the subject's waist is compared with the reference load. Newton's second law is applied to convert acceleration to GRF, using the whole-body mass in the equation. The human body mass distribution does not significantly affect the GRF, as the waist is the body's center of mass. In other words, the acceleration is multiplied by the whole-body mass.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 CV-based GRF Estimation Results

4.1.1 Data Processing

There is a self-frequency difference between the Loadsol reference (100Hz) and OpenPose (240Hz) tracking, which is intentional. For video tracking, there is a lot of noise, and this noise can be reduced by using smoothing filters. However, these filters can influence the amplitudes of walking displacement. Setting the FPS to match the Loadsol self-frequency results in the smoothed displacement losing numerous features, particularly at each amplitude. Therefore, to reduce the influence after real-time displacement signals pass through smoothing filters, set the FPS of the camera higher than the Loadsol self-frequency.

Normally, the monitored key point data cannot be directly analyzed or processed because the self-frequency of the camera is high (240 Hz in this experiment), while the walking frequencies are only between 1.5 and 2.0 Hz. Thus, it is necessary to remove some noise and high-frequency components before data analysis. The Savitzky-Golay filter is suitable for up-sampling or down-sampling the OpenPose data and enables synchronization with the reference force data (Mundt et al., 2023).

The Savitzky-Golay filter smooths and differentiates time-domain data. The idea behind this filter is to process subsets of adjacent data with least squares using a low-degree polynomial. The polynomial degree can be adjusted to achieve the desired smoothness according to the natural signal. One significant advantage of the Savitzky-Golay filter is its ability to preserve the original signal amplitudes compared to the Butterworth filter and Bandpass filter (Shahabpoor & Pavic, 2018). This preservation includes important features of the raw signal, such as amplitudes, minima, and widths, which are crucial for GRF or ground reaction load analysis.

In this experiment, the Savitzky-Golay filter is applied to smooth the displacement, velocity, and acceleration time-domain signals, while maintaining the original features and shapes of the signals.

4.1.2 OpenPose Detection Results

OpenPose can detect the subject's walking motion and locate the waist key point for each frame, allowing the 2D location of the waist key point to be recorded and exported for every frame. The key point detection results from OpenPose are exported in pixels, and only the waist displacement key point is saved for further analysis. The subject's waist displacement must be converted to real-world dimensions using a conversion factor derived from the chessboard.

The velocity of the subject's walking can be computed by taking the first derivative of the real-world waist key point displacement in the time domain. The acceleration of the subject's walking can then be computed by taking the derivative of the velocity.

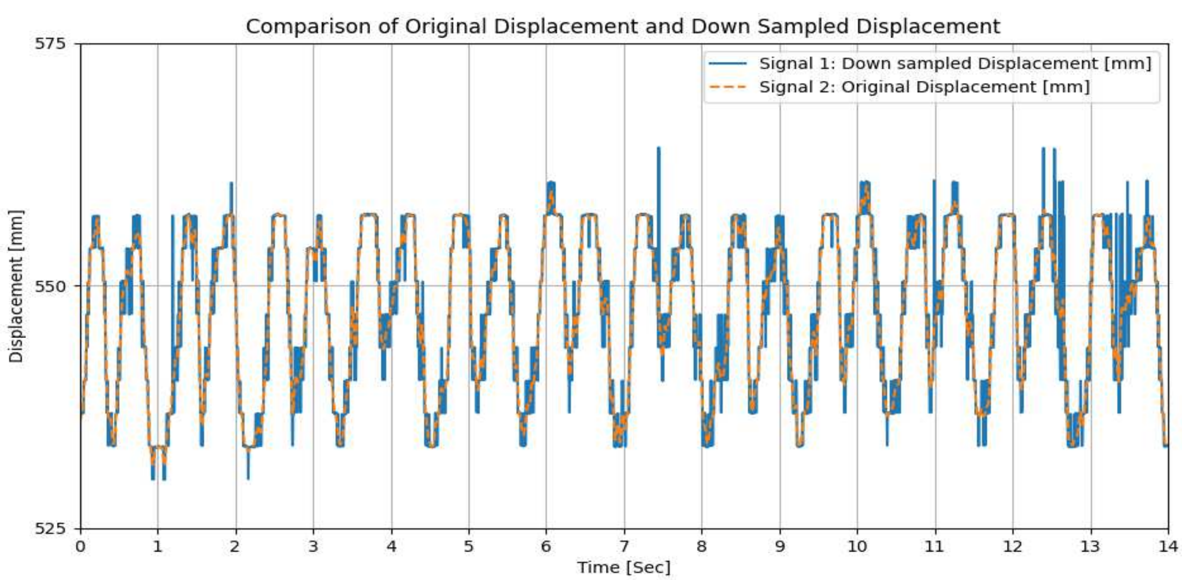

Figure 4.1. Smoothed displacement estimated by OpenPose.

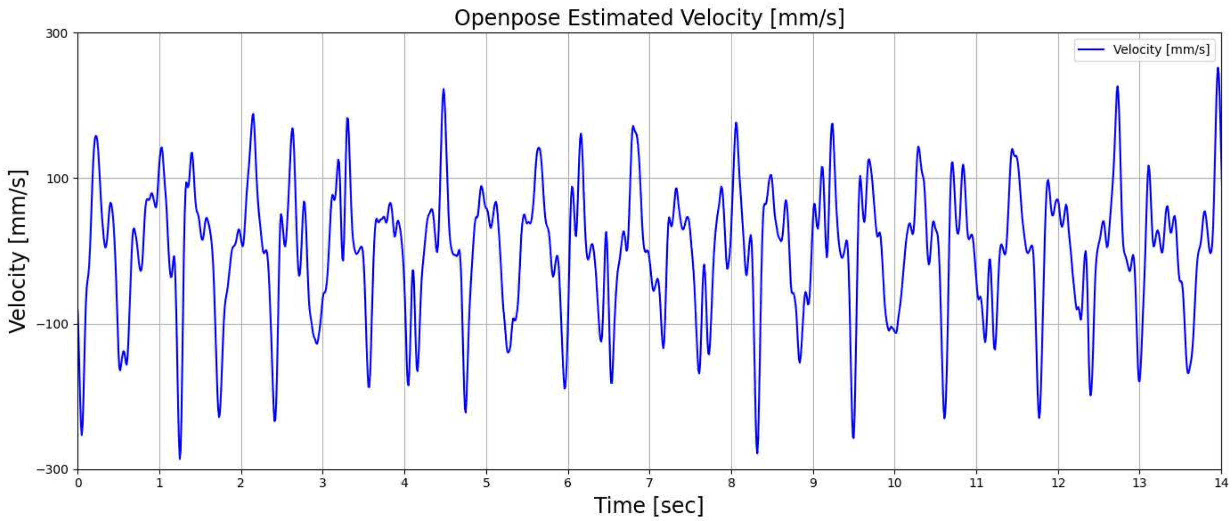

Figure 4.2. Smoothed velocity estimated by OpenPose.

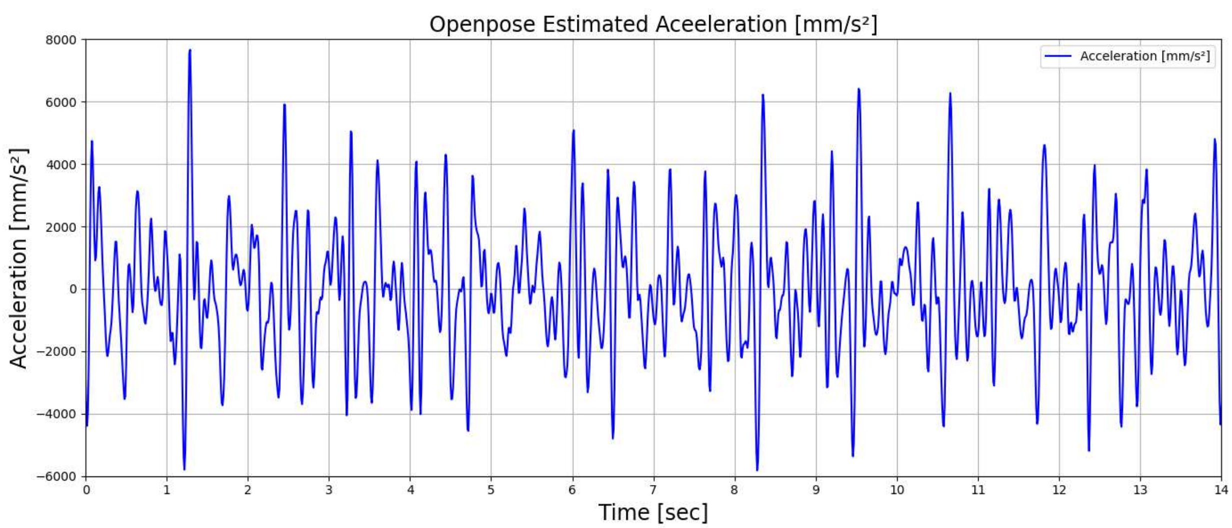

Figure 4.3. Smoothed acceleration estimated by OpenPose.

From the smoothed acceleration time-domain signal, there is still noise present, caused by the high self-frequency of the GoPro camera. As mentioned before, the self-frequency of the GoPro is 240 Hz (240 FPS), while the self-frequency of the Loadsol reference is only 100 Hz. To address this discrepancy, it is necessary to reduce the self-frequency of the GoPro-collected signal and align it with the Loadsol self-frequency. To achieve this, a Gaussian filter is applied to the smoothed acceleration time-domain data, reducing the data from 240 Hz to 100 Hz and aligning it with the reference data for further analysis.

4.1.3 OpenPose GRF Estimation

The down-sampled acceleration is then converted to GRF using Newton's second law. The results are displayed in Figure 4.4.

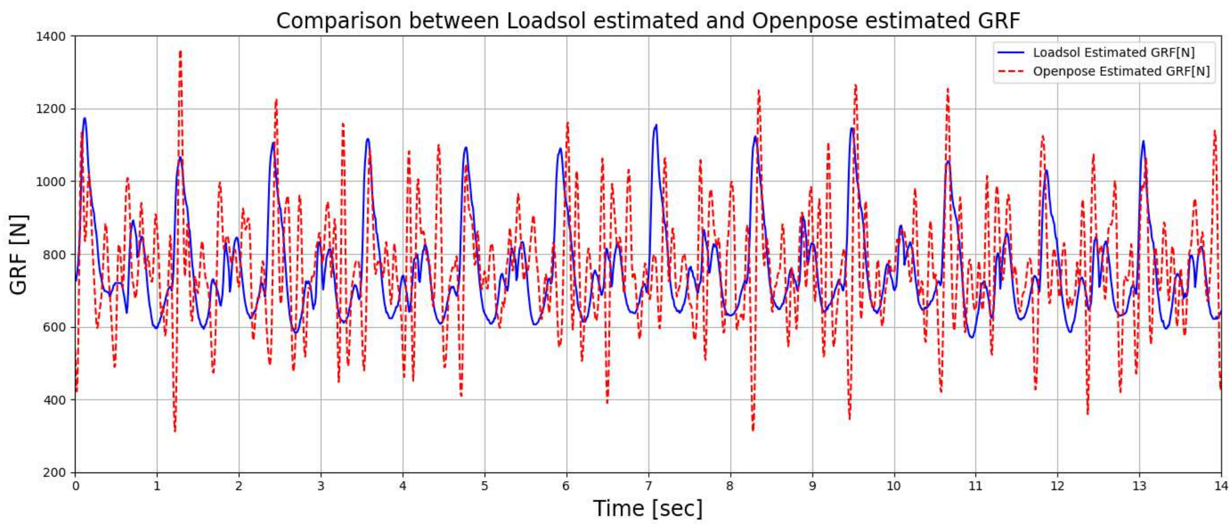

Figure 4.4. GRF collected from Loadsol and OpenPose estimation.

From Figure 4.4, the two types of measurements show similar peaks during the test, but there are high-frequency components observable in the OpenPose GRF estimation, especially between the two high curves. It is necessary to analyze the source of these high frequencies and remove them from the raw signals. In Figure 4.5, the Fast Fourier Transform (FFT) shows that the high frequencies are mainly in the range between 5 and 13 Hz compared to Loadsol power distribution. These noises are from the walking frequency and can be removed. At the walking frequency (around 1.7 Hz), the Power Spectrum Density (PSD) indicates that the two signals have similar magnitudes, which means OpenPose can estimate the GRF accurately, especially at the walking frequency. In GRF estimation, PSD helps identify the main frequencies in the forces exerted during activities like walking. By looking at the PSD, we can see how the power is spread across different frequencies and spot any patterns or noise. The rest of the extremely high-frequency components can be neglected since GRF should be in the low-frequency range only.

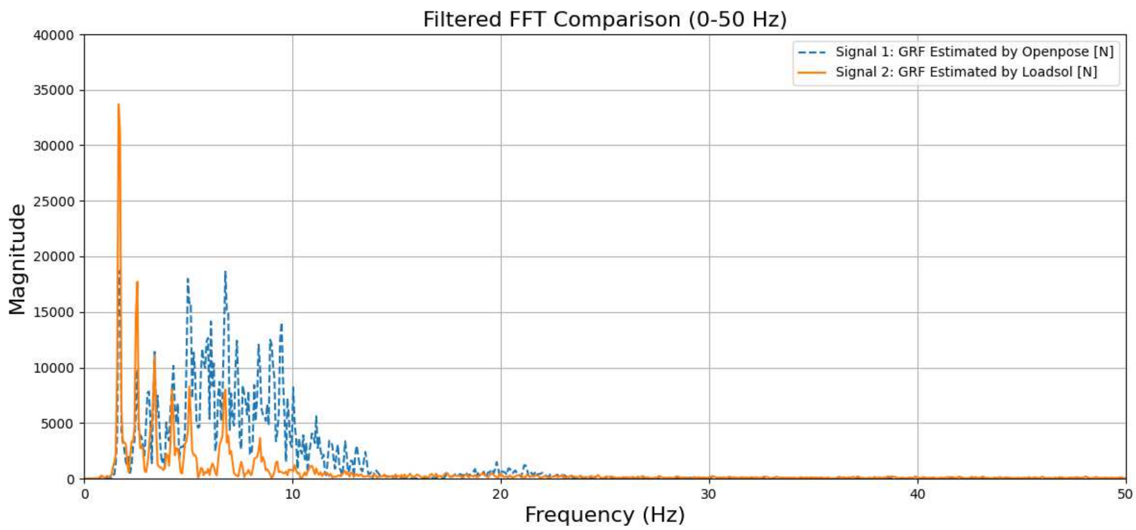

Figure 4.5. FFT for two types of estimation.

4.1.4 OpenPose Noise Identification and Removal

To analyze the origin of the high-frequency noise, a stationary test is included in this experiment. The stationary test involves the subject standing still on the treadmill without any movement, keeping the body stationary the entire time while using the same GoPro Hero 9 to record at the same FPS (240 Hz). The main goal of this test is to confirm whether OpenPose gesture estimation contains noise that affects GRF estimation. The results of the stationary displacement test are shown in Figure 4.6.

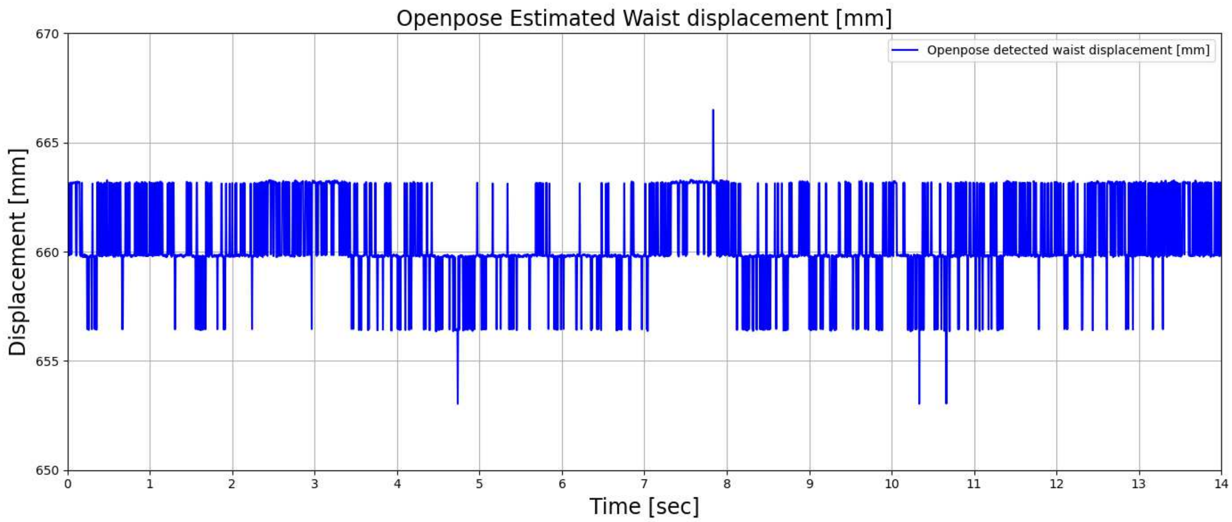

Figure 4.6. The waist key point displacement in the vertical component.

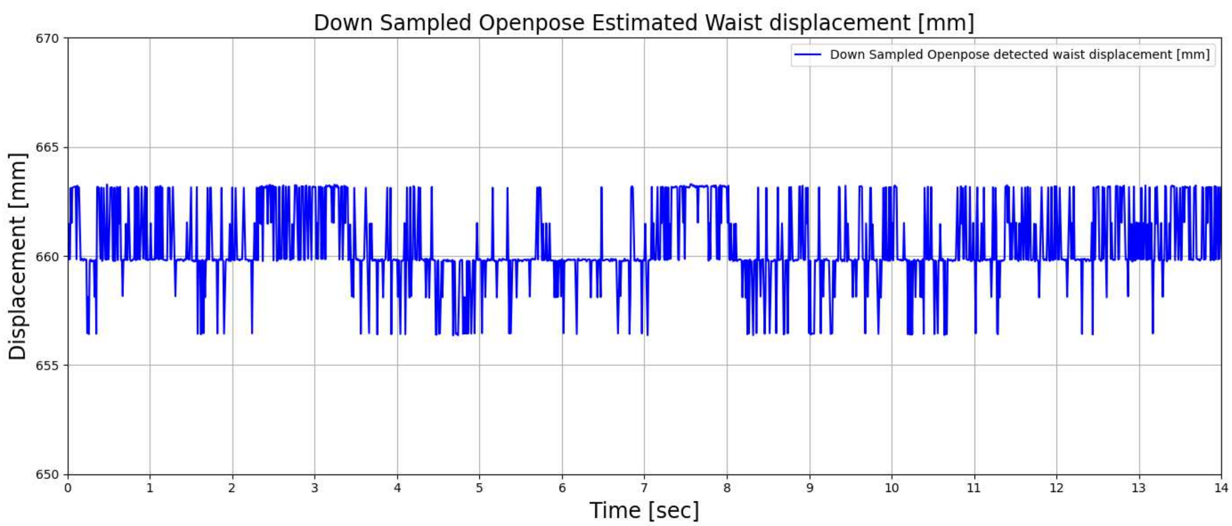

Figure 4.7. The waist key point displacement down sampled coordinates.

As observed in Figures 4.6 and 4.7, OpenPose exhibits a measurement error of about 4 mm in the vertical component during testing, constituting a tolerance of approximately 0.6% of the total test. This error is negligible, as it is extremely insignificant compared to the major vertical displacement components. However, it is important to note that these minor, rapid vertical displacements contain high-frequency information that could potentially impact further analysis, particularly in scenarios involving high-frequency self-estimation.

To address this issue, the high-frequency components must be filtered out of the signals. The initial step involves calculating the predominant frequencies of this noise. Subsequently, a bandpass filter is employed to eliminate noise within this frequency range, thereby allowing only signals within a specific range to pass through the filter. The FFT results for this noise are illustrated in Figure 4.8.

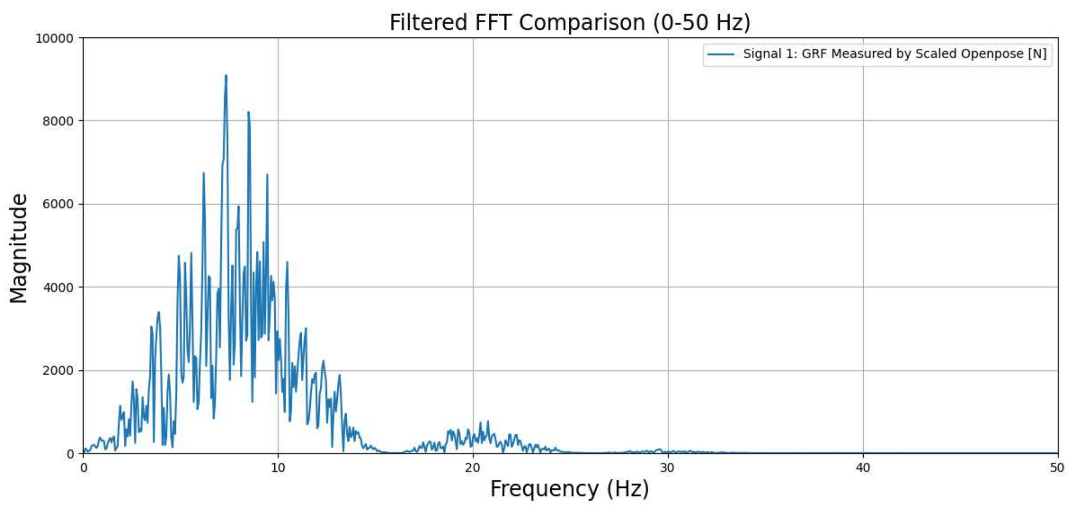

Figure 4.8. FFT for OpenPose estimation noise.

According to Figure 4.8, the main range for this noise is from 5 to 10 Hz, which is much higher than the walking frequency of interest. This means it is safe to remove all signals in the range between 5 and 10 Hz to reduce the noise, and this removal will not affect the analysis of the frequency of interest. After removing the noise, the displacement detected by OpenPose is processed with a second time derivative and converted to time-domain acceleration. Then, it is converted to GRF after applying Newton's law.

$$GRF(t) = m_{total} \times (g + a_{waist}(t)) \quad (1)$$

4.1.5 OpenPose Noise Removal Results

The estimation results from OpenPose and Loadsol are shown in Figure 4.9. Both signals indicate that the GRF ranges from 600 to 1200 Newtons. Significantly, the previous high-frequency components have been clearly removed from the original OpenPose signal, and both signals display similar fluctuating patterns in GRF. However, high-frequency components within the range of 0 to 5 Hz cannot be removed. Additionally, the peak values monitored by OpenPose are smaller than those estimated from the Loadsol reference, indicating the necessity for further signal processing.

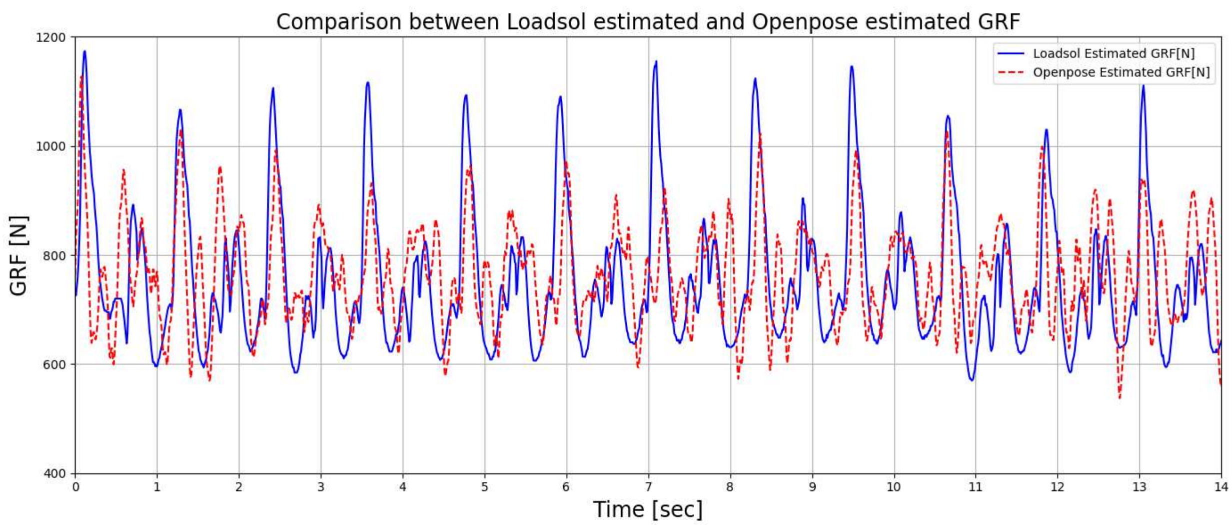

Figure 4.9. Two GRF estimated from OpenPose and Loadsol.

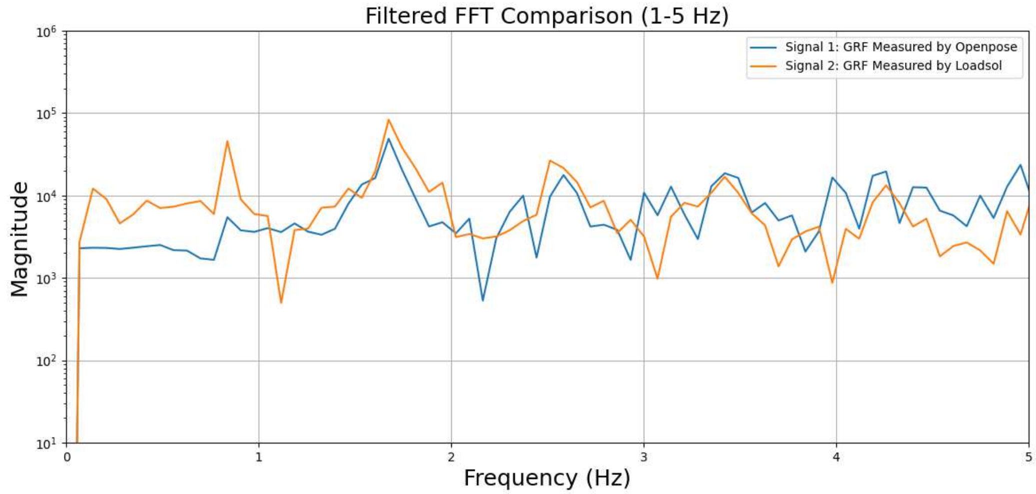

Figure 4.10. FFT for two signals in the range 0-5 HZ.

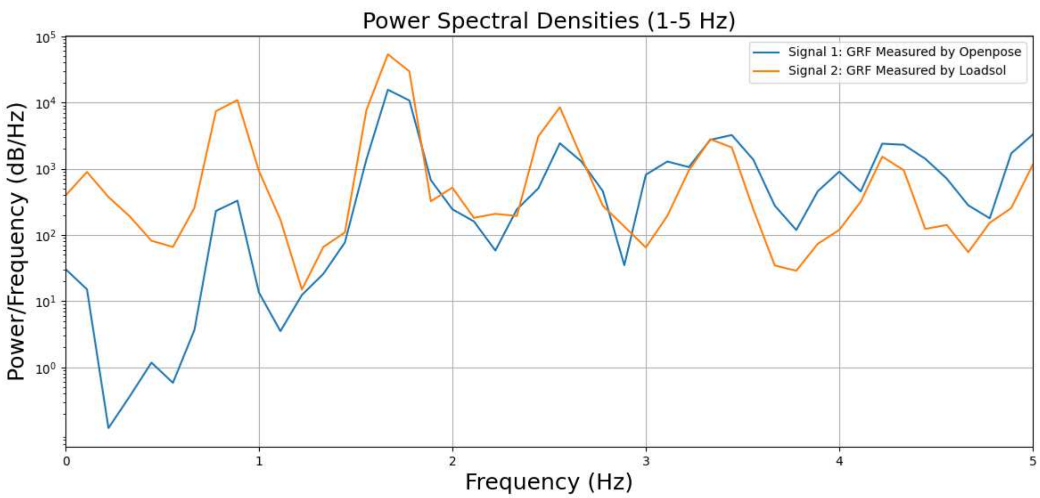

Figure 4.11. PSD for two kinds of estimation methods.

Observing the signals in Figures 4.10 and 4.11, the GRF estimated from OpenPose and Loadsol have close amplitudes and similarities in the range from 0 to 5 Hz. For the FFT results, both lines show peaks and troughs across the frequency spectrum, representing the dominant frequencies within the GRF signals.

Figure 4.10 shows the coherence values between 0 and 1, where 1 would indicate a perfect linear correlation between the two signals. This plot denotes that the two signals correspond well in the range from 0 to 5 Hz.

Figure 4.11 presents the comparison of the PSD of the two signals, indicating how the power variance of the signal is distributed over this frequency range, with most power concentrated at the GRF frequency. As a result, the two signals have peaked that line up closely, and coherence values are consistently near 1 across most frequencies, demonstrating that they are highly correlated.

As mentioned above, the signal still contains some noise after removing the main noise from the measurement error. This issue can be further addressed by applying Dynamic Time Warping (DTW) functionality. This filter relies on the Loadsol GRF estimation reference to minimize the Euclidean distance between the two signals. Without the reference GRF results, the DTW filter cannot be applied.

Shahabpoor and Pavic (2018) applied DTW to C7 human IMU estimation to average signals for different gait cycles. According to DTW theory, this algorithm finds the optimal alignment between two sequences through dynamic programming. In other words, the main goal is to minimize the total warping cost. The DTW method does not cause distortions and transformations in the time step, which enhances robustness. This property is particularly suitable for time phase variation analysis. Additionally, it is not very sensitive to noise and can deal with sequences of different time lengths. In summary, DTW can minimize the Euclidean distance between two signals, and it can also stretch or compress two sequences to allow for the best alignment between them. Finally, it adjusts temporal shifts effectively.

Therefore, DTW can be applied to this experiment to adjust the signal estimated from OpenPose. The adjusted GRF signal is shown in Figure 4.12.

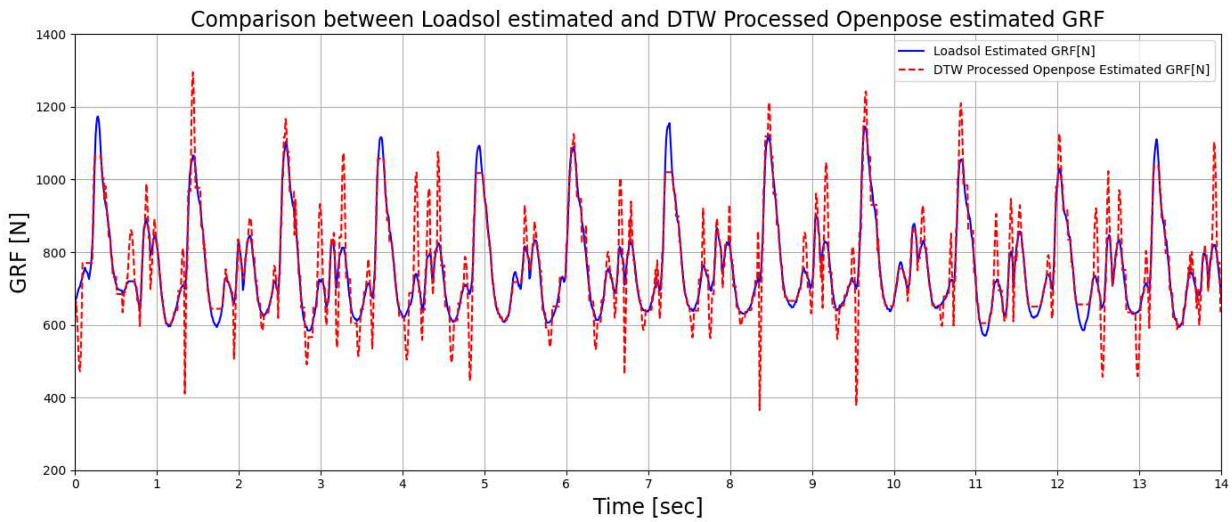

Figure 4.12. GRF after applying DTW function to OpenPose monitored signal.

Figure 4.12 reveals the absence of GRF peaks, despite the critical role of peak information in real-time GRF estimation. Therefore, the peaks collected from OpenPose need to be scaled using the Loadsol GRF reference. The results for the scaled GRF signal and reference are shown in Figure 4.13.

Figure 4.13. The comparison among scaled signal, before scaled signal, and reference GRF.

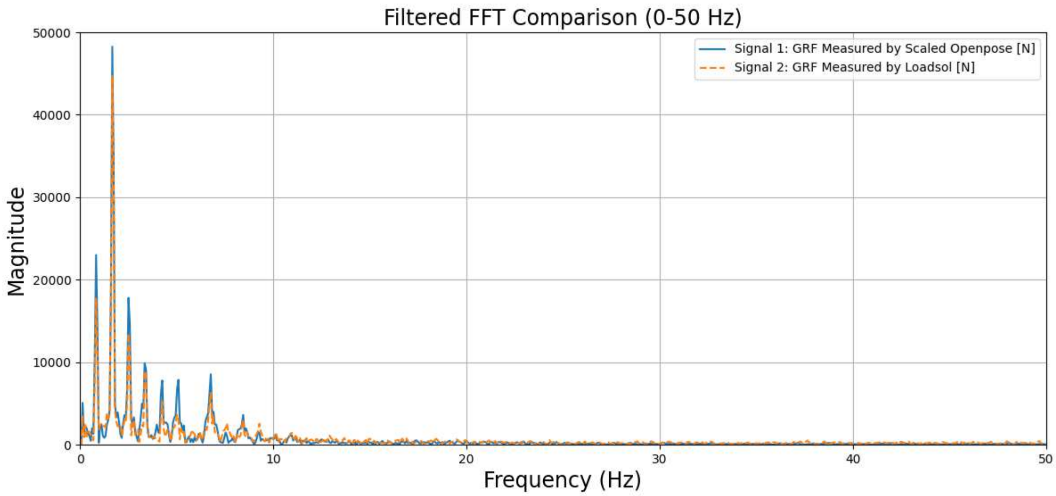

Figure 4.14. FFT for OpenPose and Loadsol GRF estimation.

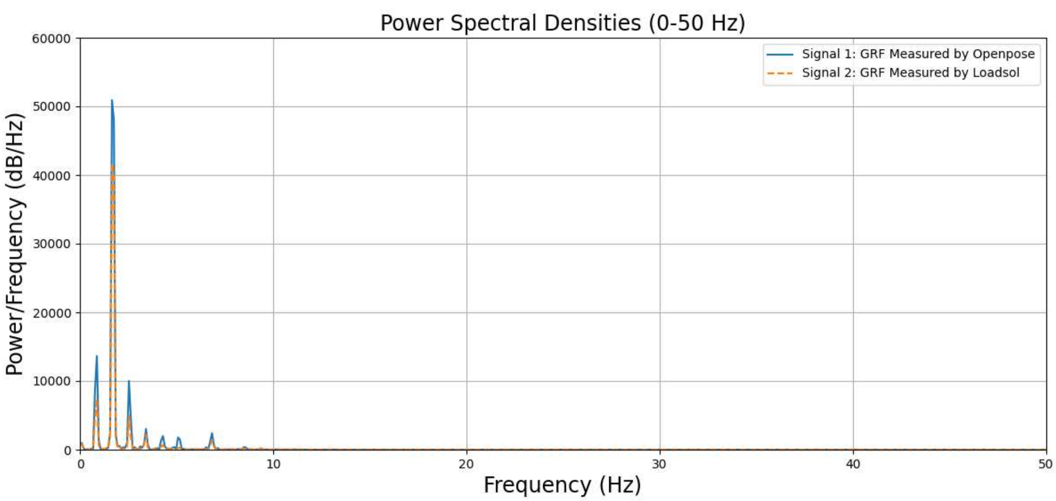

Figure 4.15. Comparison of PSD between OpenPose and Loadsol GRF estimation.

Figure 4.14 compares the FFT of the two signals. Both signals display a prominent peak at low frequencies, especially around 1.7 Hz (the GRF frequency), suggesting that the two signals have strong components around this frequency. After that, the magnitude decreases rapidly as the frequency increases. The OpenPose signal has stronger frequency components compared to the Loadsol signal, particularly at lower frequencies. The second graph shows the PSD, which measures the power present in each frequency component on a logarithmic scale, further confirming that the OpenPose signal has more power across the frequency spectrum, especially in the lower frequency range. Overall, according to these graphs, there is a strong correlation between the Loadsol and OpenPose signals at low-frequency components.

In conclusion, OpenPose is a potentially powerful GRF estimation tool because it can capture the GRF frequency and compute GRF within a tolerance of around 15% when compared with the Loadsol reference. However, there is still room for improvement in addressing the limitations of the algorithm. OpenPose can effectively measure GRF in both time and frequency domains, accurately reflecting the subject's walking frequency. However, when compared to the Loadsol GRF estimation reference, OpenPose GRF amplitudes are not perfect. Some peak GRF values are underestimated due to noise reduction filters. Additionally, the subject in this experiment had unequal walking pressure between the feet, causing adjacent peaks to differ by around 150 Newtons. Therefore, it is necessary to validate OpenPose GRF estimation accuracy by involving more volunteers for CV-based GRF estimation.

4.2 Smartphone-based GRF Estimation Results

4.2.1 Smartphone-based (MATLAB) Experiments Results

The time-domain and frequency-domain load signals collected from MATLAB and Loadsol results are shown below.

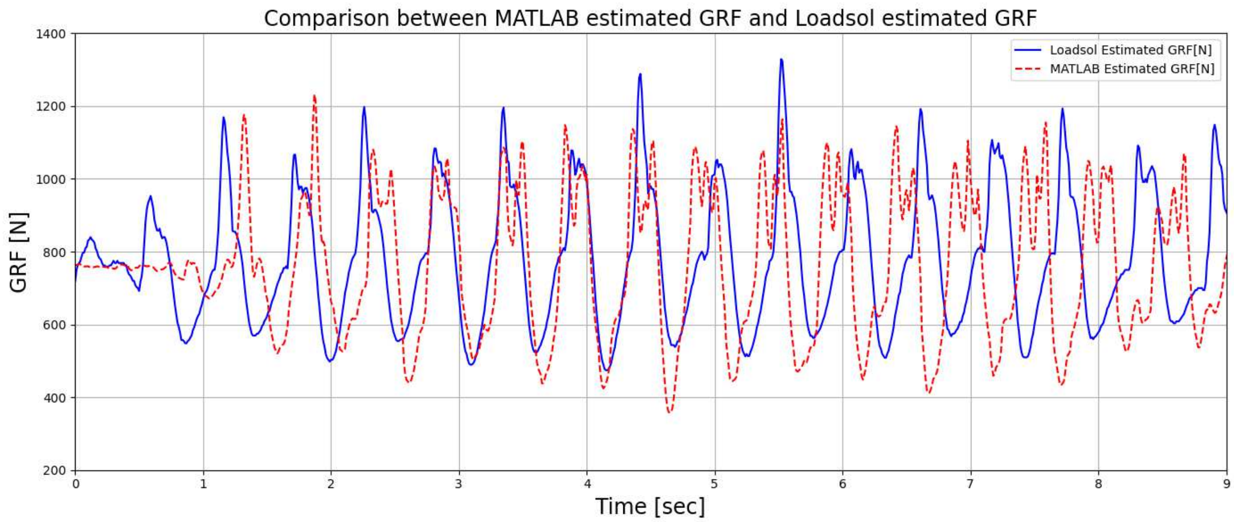

Figure 4.16. Comparison between Loadsol and phone GRF estimation.

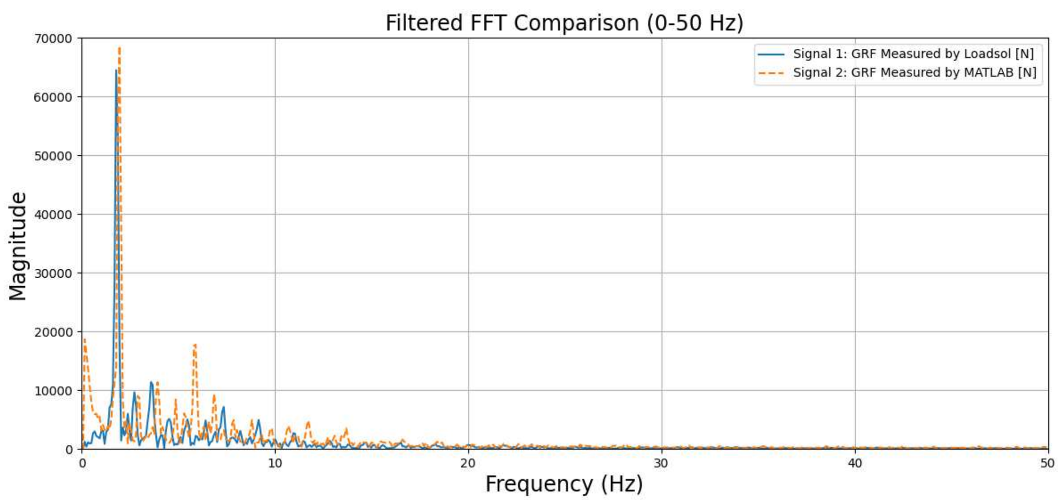

Figure 4.17. Comparison between Loadsol and phone GRF estimation after applying FFT.

According to the two figures above, the time-domain GRF collected from the waist shows a correlation with the reference GRF (Loadsol). In the frequency domain, the GRF signals also display a correlated frequency distribution. Despite the amplitudes not exactly matching, their distribution patterns still exhibit similarities.

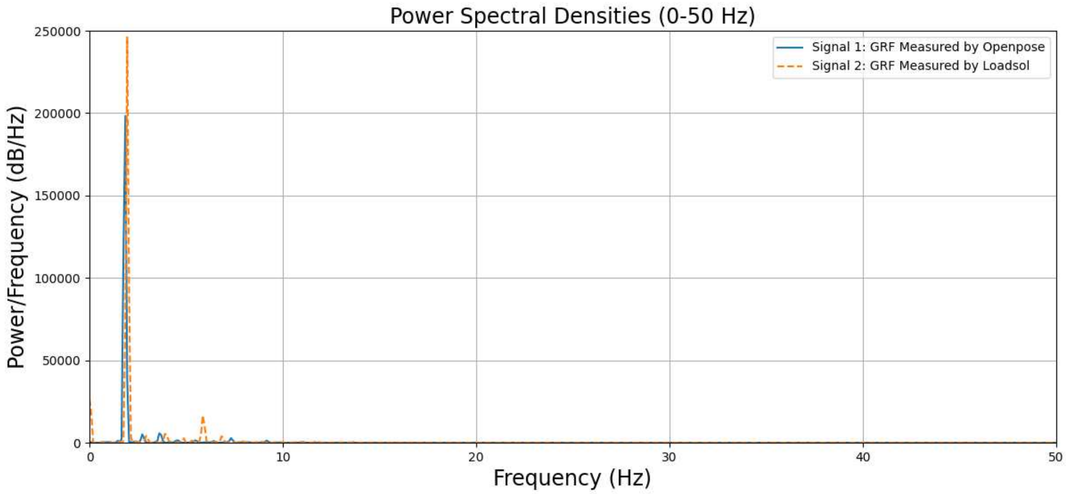

Figure 4.18. Power Spectral Densities comparison analysis for two methods of GRF estimation.

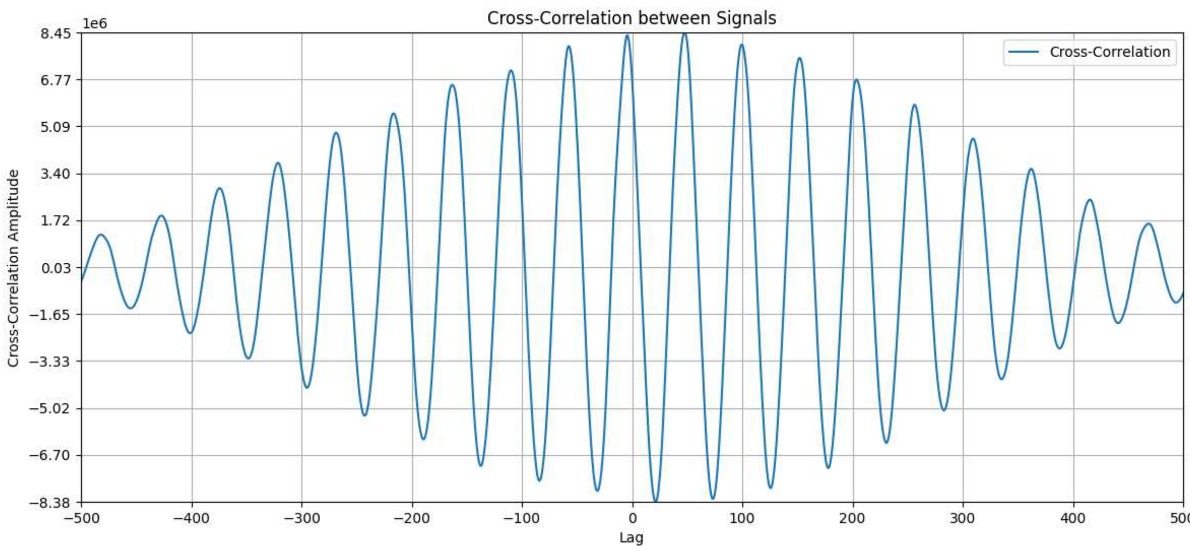

Figure 4.19. Cross-Correlation results for observing time lags between smartphone and Loadsol reference GRF estimation.

By observing Figures 4.16 and 4.19, there is a clear frequency offset between the two estimation methods, and this offset can be caused by the random offset occurring in the MATLAB Mobile App.

Cross-correlation was calculated in this experiment. As shown in Figure 4.19, the same periodic pattern indicates that the two signals are correlated at the same intervals. Zero-lag correlation indicates a strong similarity between the two signals without any shift, as the peak at zero is the highest. The amplitude of the peaks represents the strength of the correlation at different lags, suggesting a strong correlation at certain intervals. The symmetric curve typically occurs when both signals are similar in shape and share the same periodicities.

However, two secondary amplitudes indicate that the two signals do not perfectly align with each other during the test. This misalignment can be seen in the signals' plots, where the signals are offset from each other in the first and last two seconds. This misalignment is why there are two extra high magnitudes in the cross-correlation results. This issue is not unique to GRF estimation using mobile devices; it occurs in each test. These time lags can be seen as a recording issue in MATLAB mobile devices' time synchronization estimation, causing small accelerations to be recorded either slightly ahead or behind time during the test. After several tests, this issue cannot be simply solved by aligning time zones or adjusting angular degrees. Therefore, it is necessary to use alternative applications for further testing to determine whether the problem is specific to MATLAB or common to all accelerometers.

4.2.2 Smartphone-based (Phyphox) Experiments Results

4.2.2.1 Data Processing

Phyphox is introduced as an alternative application for these experiments, with the smartphones and Loadsol reference remaining the same as in the previous tests. In other words, the only variable being changed is the application system. If the random time delay issue is resolved by switching applications, it indicates a problem with the MATLAB application. However, if changing the application does not fix the issue, it suggests a potential problem with smartphone accelerometer estimation in general.

Phyphox is another mobile device physics experimental application that allows users to customize different sensors present in the mobile device for various physics experiments. These experiments include harmonic vibration, human walking, sound, light, and more.

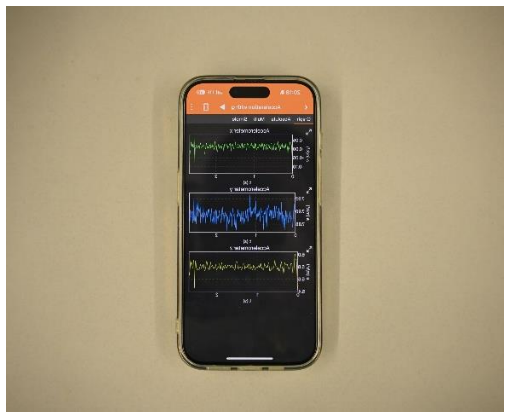

Figure 4.20. The coordinate system for Phyphox.

In the Phyphox coordinate system, the Y axis points straight upwards along the screen of the phone, the Z axis is perpendicular to the screen, and the X axis points to the right along the screen. The acceleration collected from the subject's waist and the angular velocity are exported using a laptop remote control model. Then, the pure vertical acceleration is corrected using the angular velocity in the X, Y, and Z directions as before. Finally, the pure vertical acceleration is multiplied by the center body mass and converted to GRF in units of Newtons. The results of the GRF comparison in the time domain are shown in Figure 4.21.

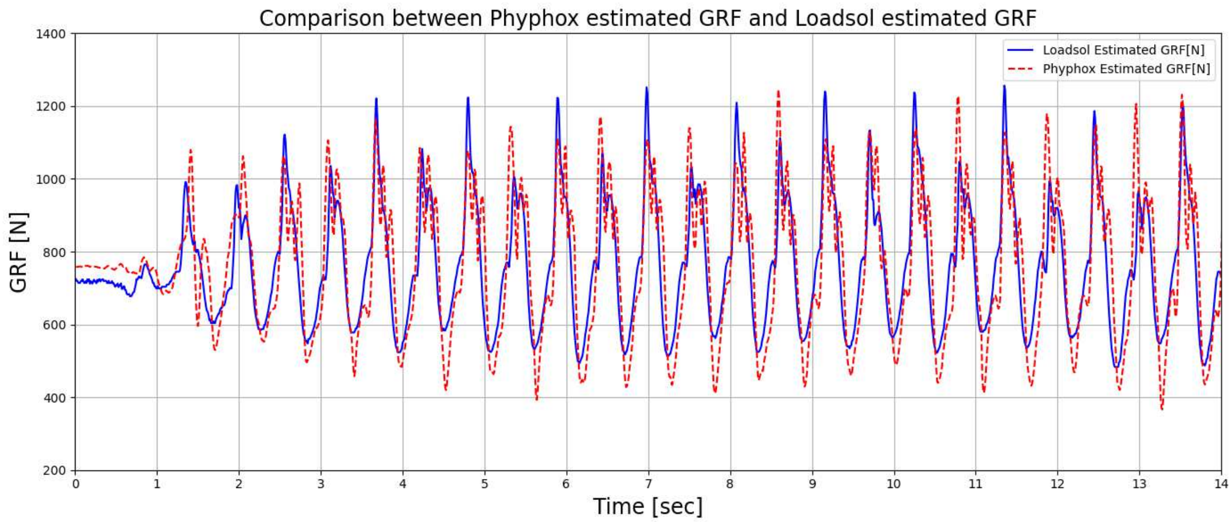

Figure 4.21. Comparison between GRF collected from Phyphox and Loadsol.

Figure 4.21, the time offset problem does not occur when using the Phyphox application. This confirms that the time offset problem is specific to the MATLAB application and does not affect all sensor applications. Additionally, the amplitudes from Phyphox estimation are significantly close to the reference load amplitudes.

There is a special observation from this test: the first step GRF from the reference Loadsol and Phyphox estimation are significantly different. This amplitude difference is much larger than for other steps and is a common issue observed in every test. The mobile device estimation load is usually much larger than the reference GRF for the first two steps. This issue is caused by Loadsol error estimation for the first two steps, as observed from Figures 4.22 to 4.23. At the beginning of the first two steps, there are two smaller amplitudes (780 Newtons and 800 Newtons, respectively) from the left Loadsol and right Loadsol, and these smaller amplitudes cause a large difference in the first two steps of GRF estimation.

The reason for this discrepancy is that Loadsol estimation is based on pressure integration from multiple sensors within the Loadsol device. The Loadsol device integrates the pressure to calculate the load. However, this pressure integration has strict requirements for the walking position. For example, if only the heel, half side of the foot, or toes are in contact with the Loadsol, it can cause integration errors. This situation typically occurs during the first few steps of human walking because the subject usually needs to adjust their walking speed and movement. As a result, only the toes, half side of the feet, or heels might be in contact with the Loadsol during this initial period.

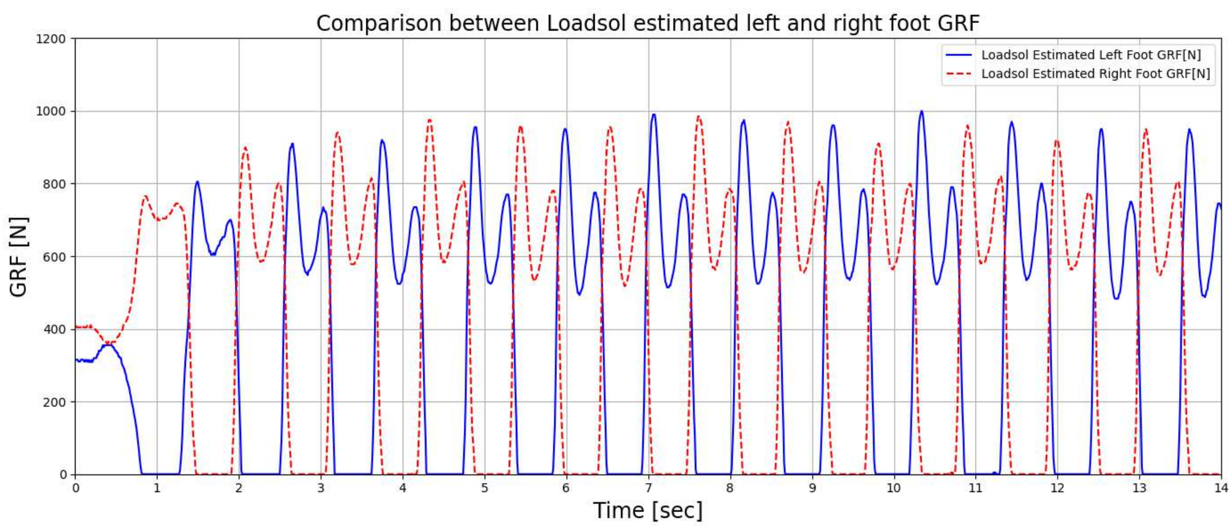

Figure 4.22. GRF collected from left Loadsol and right Loadsol.

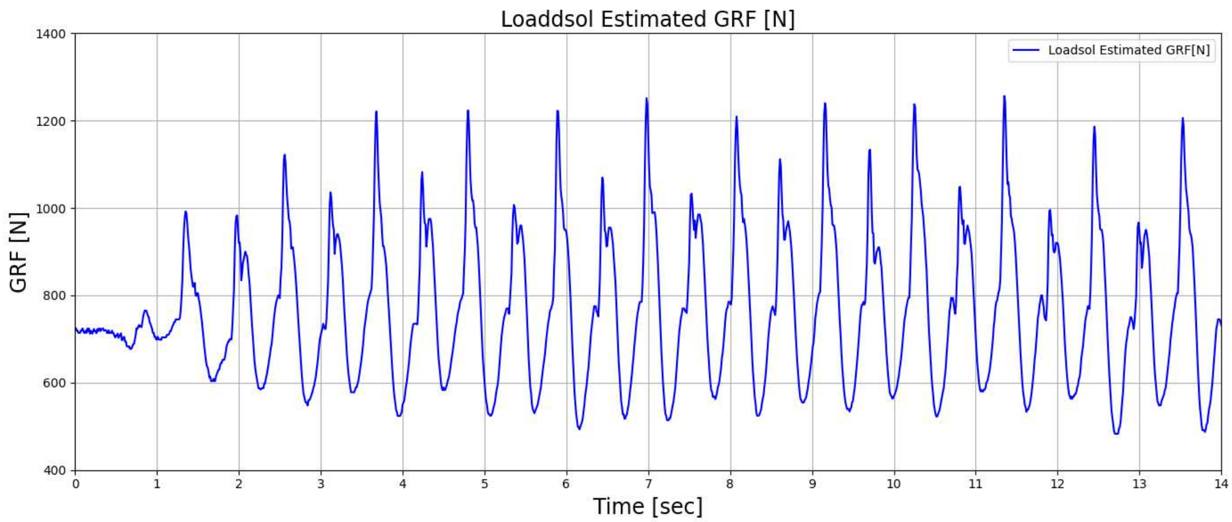

Figure 4.23. Total GRF estimated by Loadsol.

From observations in the previous figures, there is a constant load between 0 and 1.5 seconds. This constant period is caused by the subject's reaction after hearing the start command. Therefore, this period should be removed from the data processing.

4.2.2.2 Phyphox GRF Estimation Results

The removal of reaction signals is displayed in Figure 4.24.

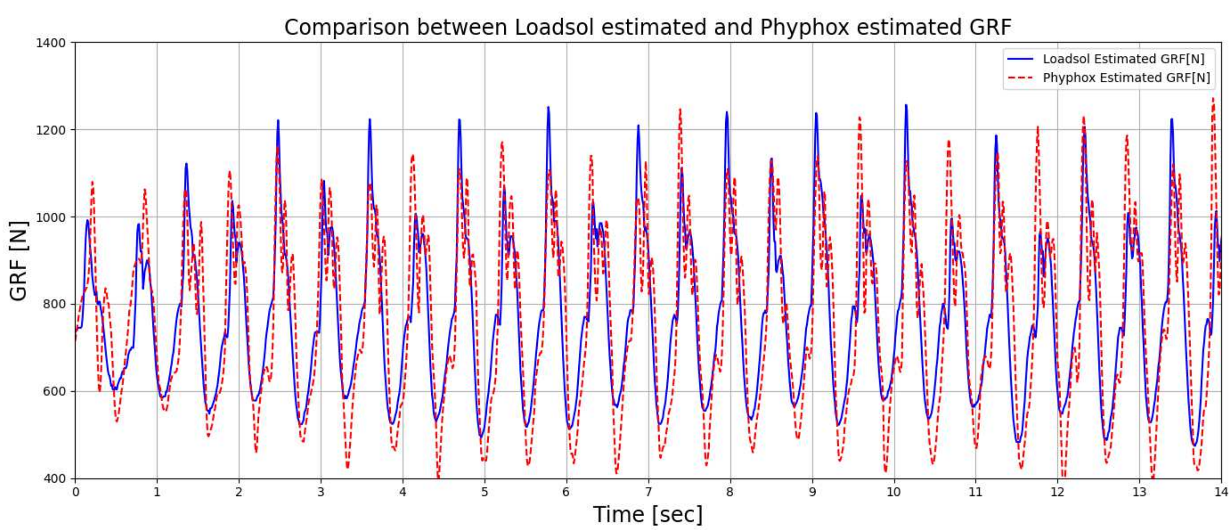

Figure 4.24. GRF comparison after removing reaction period.

The processed signals are transformed using FFT and MSC. The results are shown in Figures 4.25 to 4.27. This process aims to analyze the amplitude features at walking frequency and determine whether the mobile device accelerometer GRF estimation is as reliable as the Loadsol reference, especially at walking frequency.

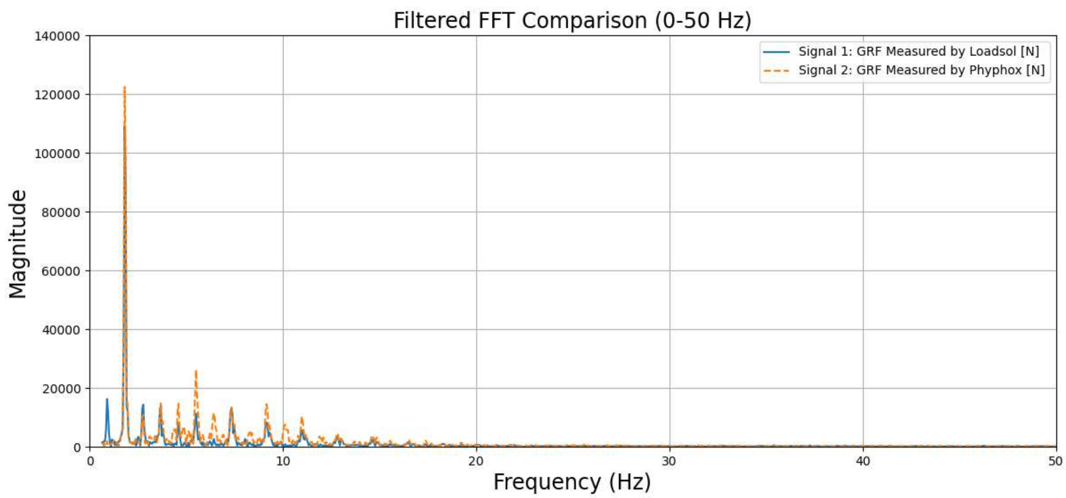

Figure 4.25. FFT for two signals collected from Phyphox and Loadsol.

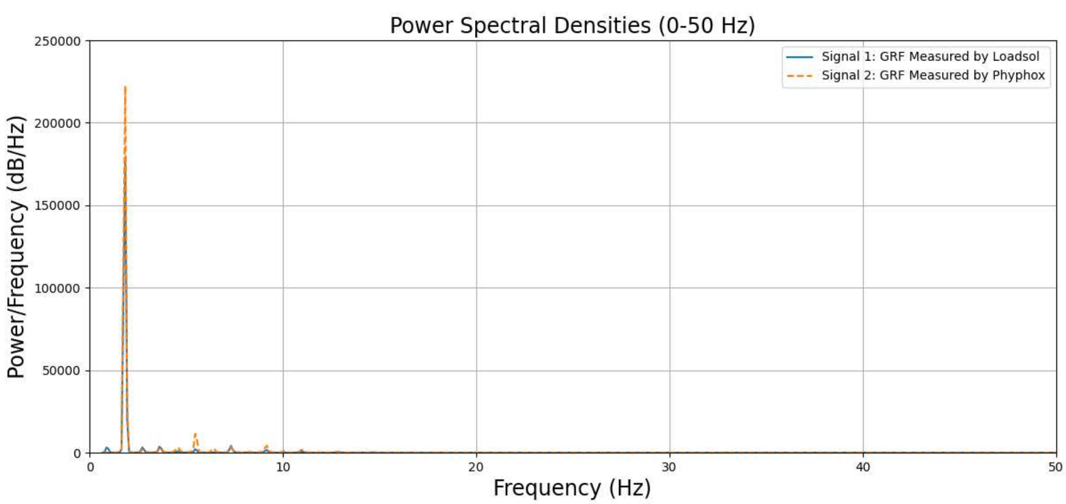

Figure 4.26. Comparison of PSD between two signals collected from Phyphox and Loadsol.

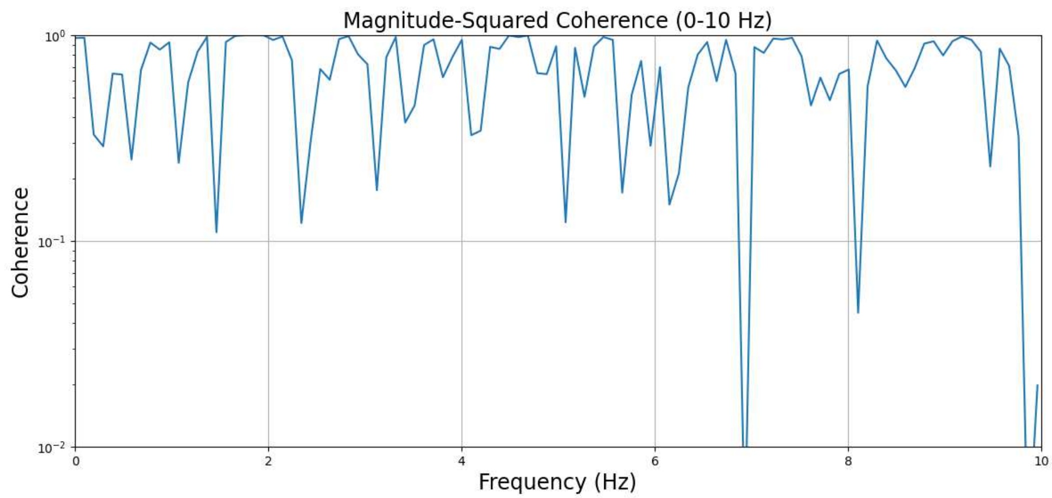

Figure 4.27. Magnitude Squared Coherence between two signals.

From the FFT analysis, it is evident that the most matched frequency period is from 1.5 to 2.1 Hz, which corresponds to the subject's comfortable walking frequency range. In this range, the magnitude of the signals is highly correlated with each other, indicating that mobile devices can estimate GRF well within the walking frequency range.

By observing the PSD diagram, the power density exported by mobile devices closely matches the Loadsol collected signal. This means that the walking power density perfectly coincides with the reference signal in the range from 1.5 to 2.1 Hz.

In the Magnitude Squared Coherence (MSC) diagram, which is a statistical evaluation used in signal processing, the degree of linear correlation between the CV-estimated GRF and the Loadsol reference GRF in the frequency domain is determined. Specifically, MSC quantifies how well one signal can be predicted from another in the frequency domain. MSC results confirmed that in the range of 1.5 to 2.1 Hz, the coherence is close to 1. This indicates that the two signals collected from the mobile devices and Loadsol are linearly related at this frequency range. In other words, mobile devices accurately estimate GRF at the walking frequency range and exhibit almost perfect correlation with the Loadsol reference.

From the observations above, all the signal data analysis and comparison strongly indicate that the signal collected from mobile devices is highly correlated with the Loadsol reference signal, especially for magnitudes in the range from 1.5 to 2.1 Hz, which is the subject's comfortable walking frequency range. Therefore, this strong correlation between the two signals justifies that mobile devices are not only sufficient for estimating the GRF but, more importantly, can also address the tracking error in the first two steps of the Loadsol reference by collecting acceleration from the waist of the subject.

4.2.2.3 Comparison Between Phyphox and MATLAB GRF Estimation Accuracy

There is a comparison of GRF estimation results from MATLAB and Phyphox (Table 4.1).

Table 4.1. The correlation coefficients from MATLAB and Phyphox measurements.

	RMSE	Pearson Correlation Coefficient	Euclidean Distance	Standard Deviation
Phyphox	115.74	0.83	4983.03	180.92
MATLAB	245.59	0.13	7509.86	179.99

The Root Mean Square Error (RMSE) measures the average size of errors between predicted values (smartphone measurements) and reference values (Loadsol estimates). The value range can influence the RMSE, which gives more weight to large errors Therefore, we also consider the standard Deviation (SD) to account for data variability. If the RMSE is less than the SD, the error is considered relatively small.

As shown in Table 4.1, the RMSE for the Phyphox estimate is 115.47, which is less than the SD of 180.92. This means the error for the Phyphox ground reaction force (GRF) estimate is small. However, the RMSE for the MATLAB estimate is 245.59, much higher than the SD of 179.99. This larger error is due to a time offset issue.

The Pearson Correlation Coefficient (PCC) measures the linear correlation between the measured GRF and the Loadsol reference. The PCC for Phyphox is 0.83, which is higher than MATLAB's 0.13, indicating a better correlation for Phyphox. The Euclidean Distance measures the difference between the measured GRF and the reference, with a larger value indicating a larger error. Based on the values from Phyphox and MATLAB, Phyphox provides a better GRF measurement compared to MATLAB.

In summary, MATLAB has a time step offset problem for real-time GRF estimation, while Phyphox estimates GRF well in both correlation and RMSE results. Therefore, Phyphox will be used for further GRF estimation in this thesis.

4.2.2.4 DTW Processed Phyphox GRF Estimation

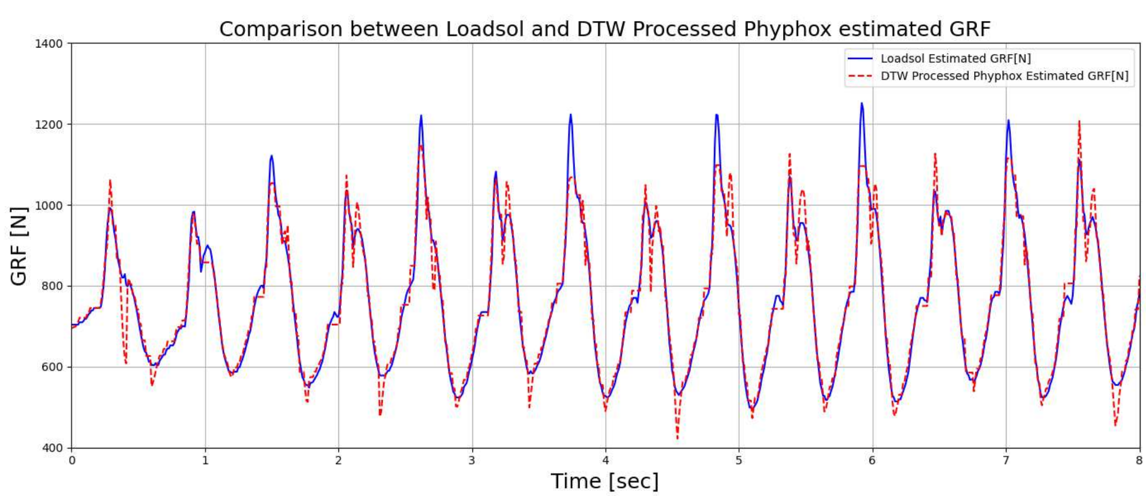

Figure 4.28. Comparison of reference signal and DTW processed signal.

Finally, the GRF collected from Phyphox is processed using DTW, which minimizes the Euclidean distance between the raw signal and the reference signal. The normal offset mathematical method cannot be applied to this problem. Instead, DTW can solve this problem by using temporal sequences that can vary in speed. For instance, DTW can detect similarities during walking, regardless of differences in walking speeds or variations in accelerations and decelerations during observation.

The comparison between the reference signal and the adjusted signal is plotted in Figure 4.29. The two signals exhibit perfect time synchronization and closely match amplitudes for each gait.

In conclusion, for mobile device applications, MATLAB has a significant time recording problem that cannot be solved by aligning the time zone. This error causes serious delays and advances during the test, leading to a complete walking estimation offset, which is unacceptable for real-time ground reaction force estimation. However, Phyphox does not have this issue and can estimate GRF accurately compared to the Loadsol reference. Therefore, Phyphox will be used for further GRF estimation analysis.

4.3 Comparisons

4.3.1.1 Objectives Information

In this chapter, the serviceability of OpenPose and smartphones for GRF estimation is compared, and the results confirm that smartphone GRF estimation performs better than OpenPose estimation. The OpenPose GRF estimation method is tested on four volunteers, and the reliability of OpenPose estimation will be discussed.

In the previous experiments, all CV-tests were conducted on the same person (objective 1). However, the results cannot be confirmed by only one subject, as many parameters can affect the experiment, such as body weight, comfortable walking frequency, and walking motion. Therefore, it is necessary to conduct additional analyses by testing the same process on different subjects while keeping the parameters for Loadsol and Phyphox the same. In this alternative experiment, four more subjects are involved, and their information is shown in Table 4.2 below.

The volunteers, all female, range in height from 163 cm to 171 cm and self-weight from 65 to 70 kg. They walk at their comfortable frequency of 1.6 to 1.7 Hz during the test. The photos taken during the tests and the OpenPose detection results are shown in Figure 4.31. During the test, the five volunteers are asked to walk on the treadmill at their comfortable walking frequency and maintain that frequency. They are informed that their walking is being recorded, and their smartphones are fixed to their waists (around the navel by using tape wrapped around the waist). The data is then exported for further analysis. After the same signal processing as previously, the signal collected from the smartphone is corrected to vertical GRF, and the signal exported from OpenPose is processed by down-sampling and noise removal.

Table 4.2. The Alternative subjects' data collection table.

	Subject 2	Subject 3	Subject 4	Subject 5
Height	171 cm	163 cm	168 cm	170 cm
Body Weight	63 Kg	70 Kg	62 Kg	63 Kg
Loadsol Size	V (8-9 US Size)	V (8-9 US Size)	V (8-9 US Size)	V (8-9 US Size)
Number of Sensors	99	99	99	99
Pressure Range	20-600	20-600	20-600	20-600
Loadsol Self-frequency	100 Hz	100 Hz	100 Hz	100 Hz
APP	Phyphox	Phyphox	Phyphox	Phyphox
APP Self-frequency	100 Hz	100 Hz	100 Hz	100 Hz
Walking Frequency	1.6 Hz	1.6 Hz	1.7 Hz	1.7 Hz

4.3.1.2 OpenPose Detection Results Comparison

This process aims to clarify that mobile device accelerometers and OpenPose can estimate GRF accurately under varying subjects' parameters. Once this analysis is complete, the reliability of mobile device accelerometers and OpenPose estimation can be officially confirmed.

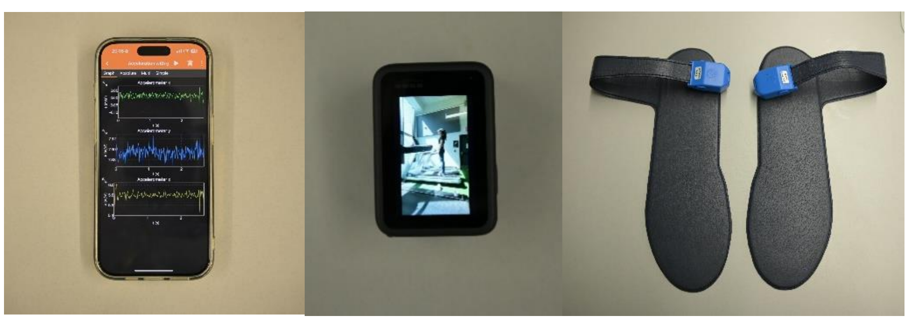

Figure 4.29. Phyphox, GoPro, and Loadsol.

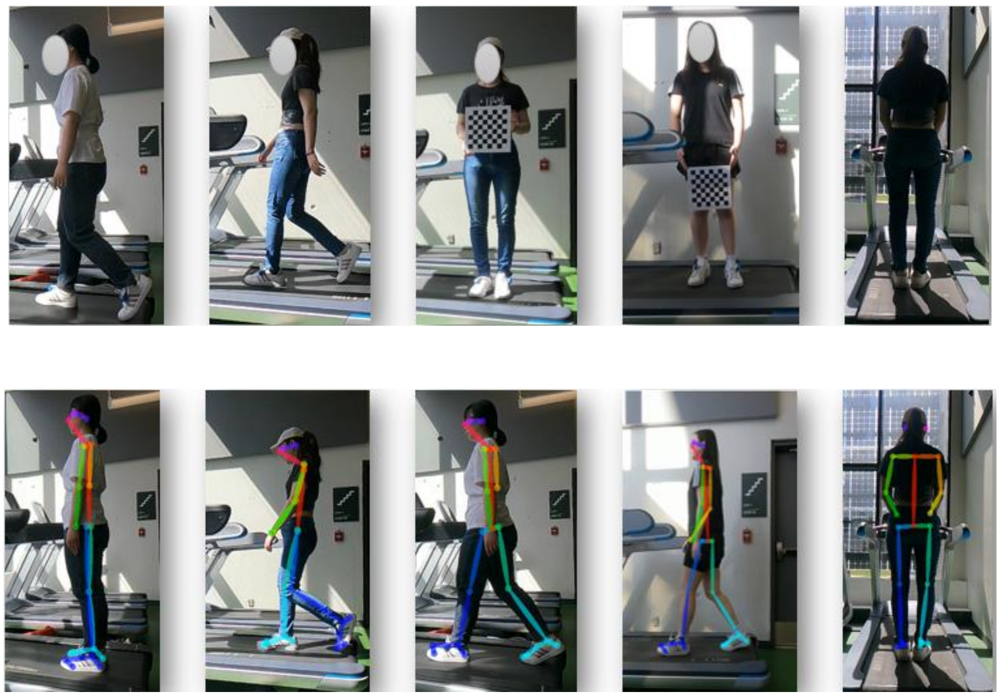

Figure 4.30. The photos during test and OpenPose detection results.

The results from OpenPose GRF estimation for all the subjects are shown in Figures 4.31 to 4.38.

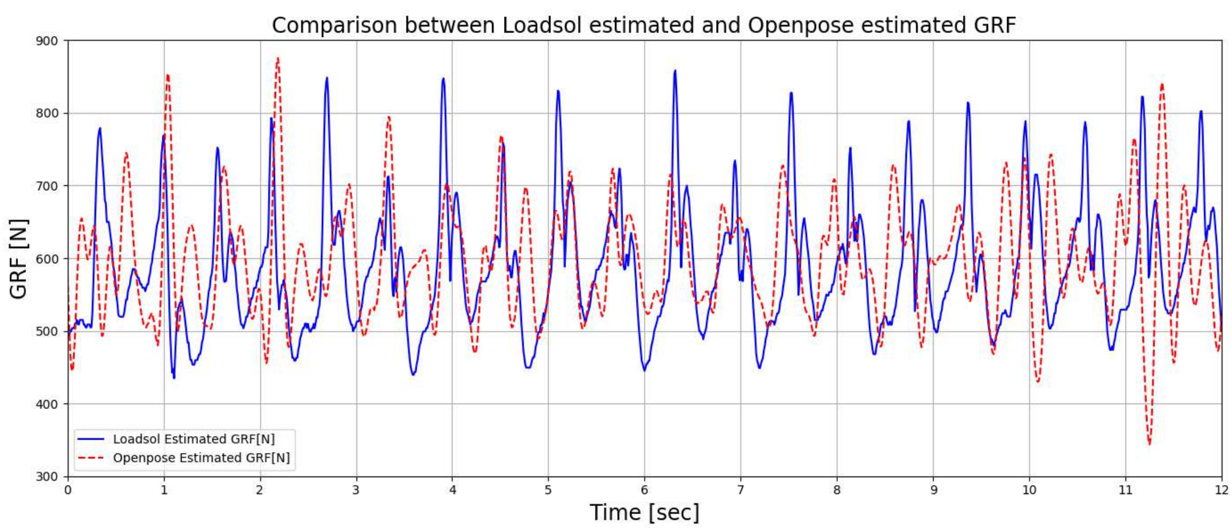

Figure 4.31. The comparison between GRF measured from Phone and OpenPose.

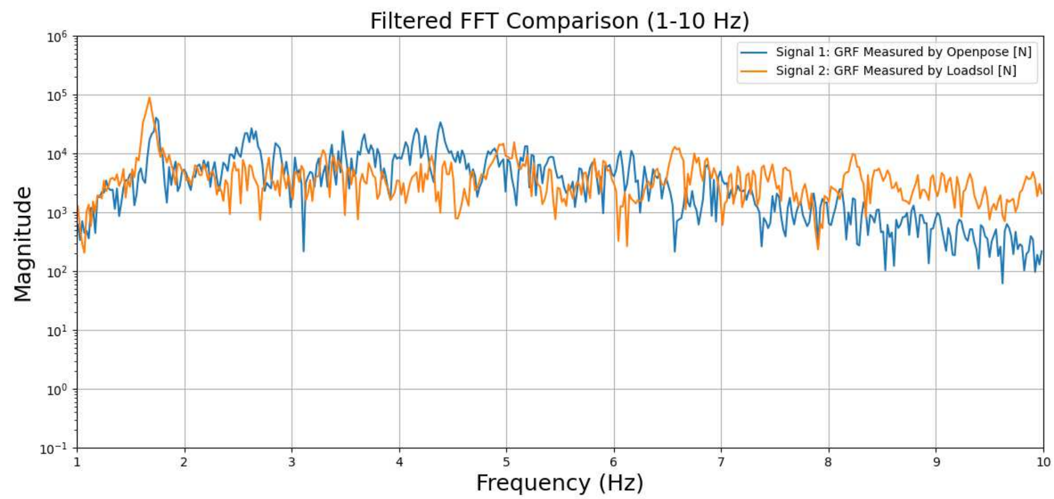

Figure 4.32. The comparison between FFT processed GRF measured from Phone and OpenPose.

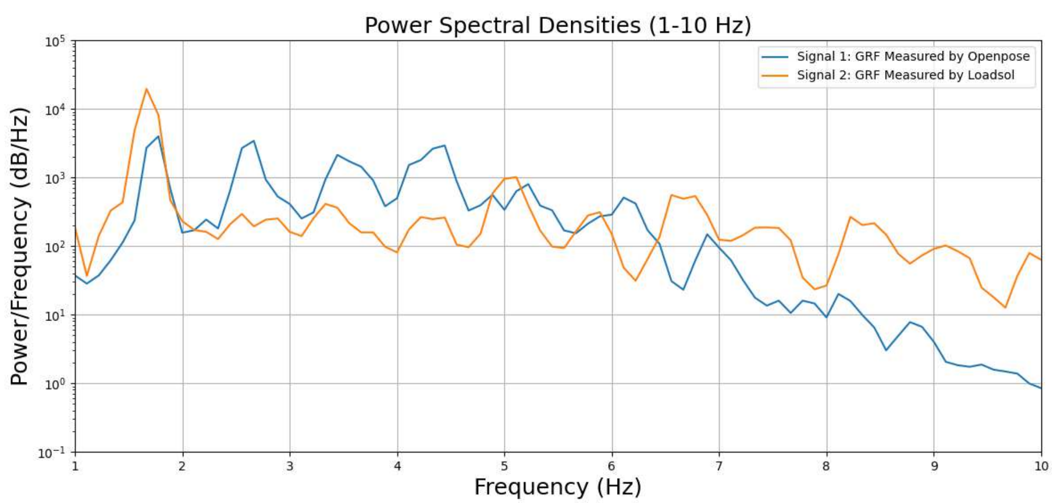

Figure 4.33. The comparison of PSD between phone and OpenPose measurement.

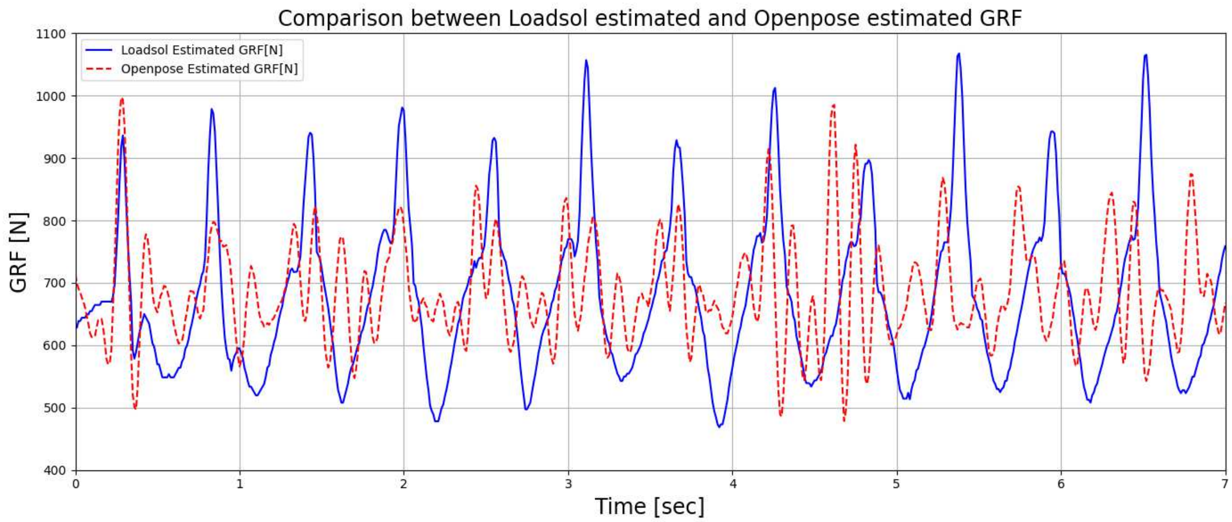

Figure 4.34. The comparison of GRF measured by OpenPose and Phone for the second objective.

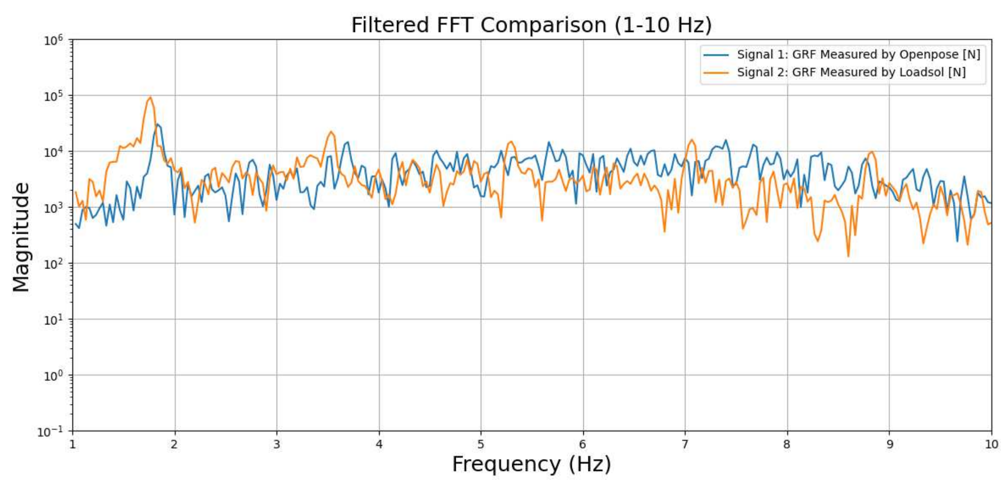

Figure 4.35. The comparison of FFT between signal collected from OpenPose and Phone for the second objective.

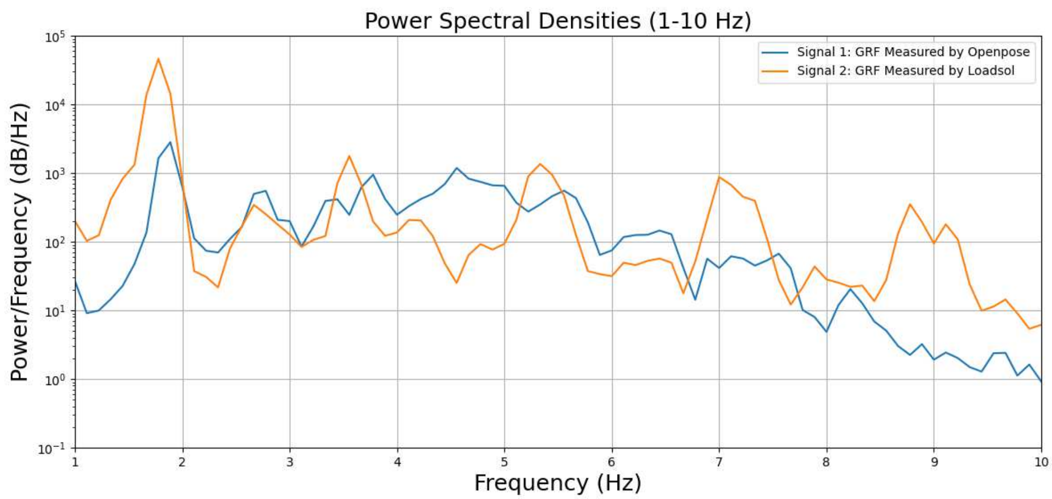

Figure 4.36. The comparison of PSD from OpenPose and Phone GRF collection for the second objective.

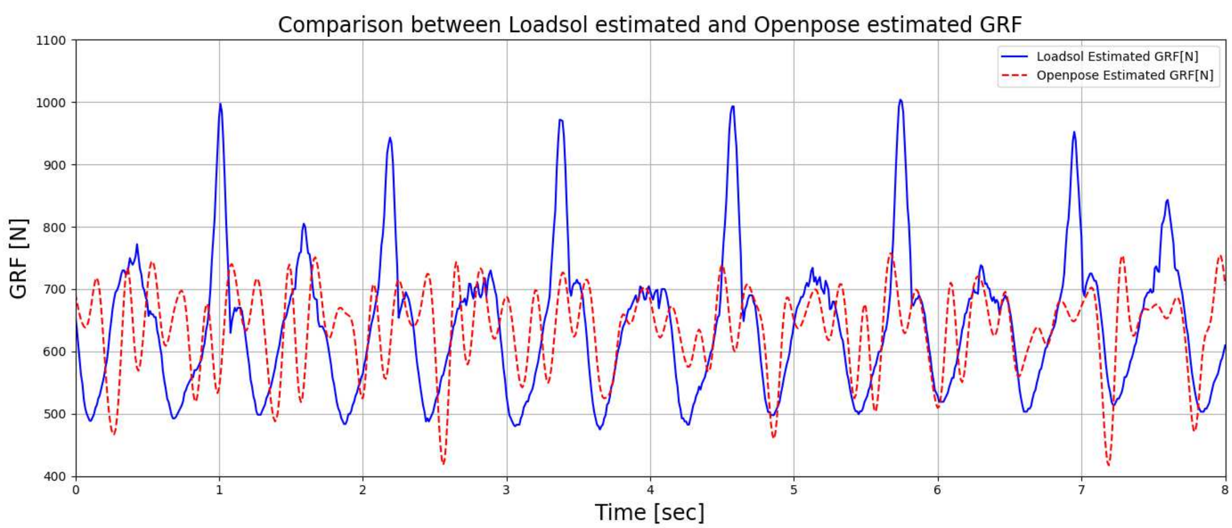

Figure 4.37. Comparison between OpenPose and smartphone GRF estimation for the third objective.

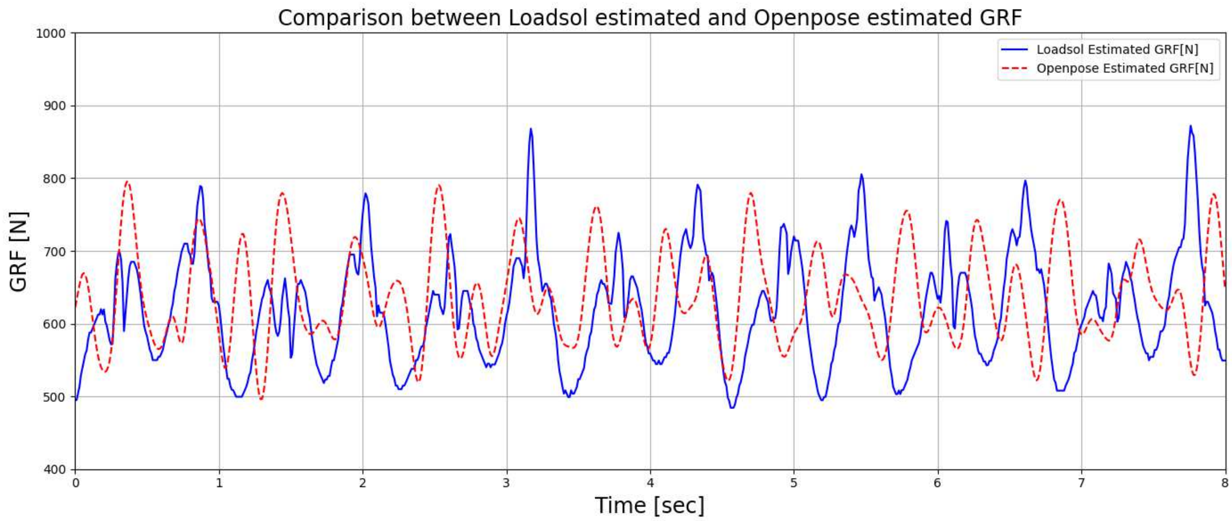

Figure 4.38. Comparison between OpenPose and smartphone GRF estimation for the fourth objective.

Table 4.3. The objects' alternative impact information for OpenPose GRF Estimation.

	Object 2	Object 3	Object 4	Object 5
Height	170 cm	163 cm	168 cm	170 cm
Weight	62 Kg	70 Kg	65 Kg	65 Kg
RMSE	104.98	134.93	116.15	86.80
Standard Deviation of OpenPose	72.51	82.96	68.71	67.53
Standard Deviation of Loadsol	78.96	127.28	110.35	72.43

As seen in the previous figures from 4.31 to 4.38, the magnitude of GRF collected from the Loadsol is slightly higher than the GRF computed by OpenPose. This issue also surfaced in Chapter 3's CV-based GRF estimation and it continues to be a common concern in this research. Additionally, by observing the comparison between OpenPose and smartphone estimated GRF, another issue appears during the alternative tests: a slight time offset.

This time series offset did not occur during the phone and Loadsol comparison because the time series were controlled by the local time zone. However, the GoPro time series cannot be controlled by the local time zone and are instead controlled by the frame rate. After down-sampling and applying smoothing filters, the time series can be offset from the reference signal. Nevertheless, this offset is not significant compared to the reference, with most offsets around 0.5 seconds for each peak.

From the following FFT and PSD analysis, the magnitude of signals at the walking frequency (1.6 Hz) is close to the Loadsol reference amplitudes, indicating that OpenPose can measure the key features of GRF during the test. It shows a similar decreasing trend in FFT and PSD processing. In other words, OpenPose can measure the GRF when the Loadsol is serving as a reference on the treadmill and is not affected by alternative factors such as height, weight, or walking frequency.

However, the underestimation of GRF is a primary issue in this research. This underestimation is mainly caused by the smoothing and down-sampling filters used during noise removal processing, which reduce the amplitudes of OpenPose-detected GRF. The main noise comes from OpenPose skeleton key point detection. The video is recorded at 240 FPS during the test, meaning there are 240 frames per second, and OpenPose needs to detect key points for over 9000 frames during the recording. Consequently, the key point detection at the waist varies over time, causing unstable displacement detection that needs to be removed during signal processing.

This issue could potentially be solved by upgrading the COCO dataset of OpenPose or upgrading the model itself, which can improve the accuracy of detection during high-quality and high-speed video recording. If this issue can be resolved through COCO data training, then smoothing and down-sampling will not be necessary in the future, preserving the original amplitudes after processing. By observing from 5 subjects, there is a special finding that subject 3, who is much shorter and heavier than the other four subjects, got the highest RMSE results (12%) and the most non-ideal estimation accuracy. This might be caused by the lack of variance of COCO training data, which, as a result, caused the low accuracy of GRF estimation.

In addition to analyzing the reliability of smartphones and OpenPose detection, the recording position is also examined in this study. Four volunteers were recorded from the side and back, respectively, and the results of GRF estimation are compared to the Loadsol reference. The results indicate that recording from the side yields much better results than recording from the back, especially in terms of waist key point detection accuracy.

4.3.1.3 Smartphone Estimation Results Comparison

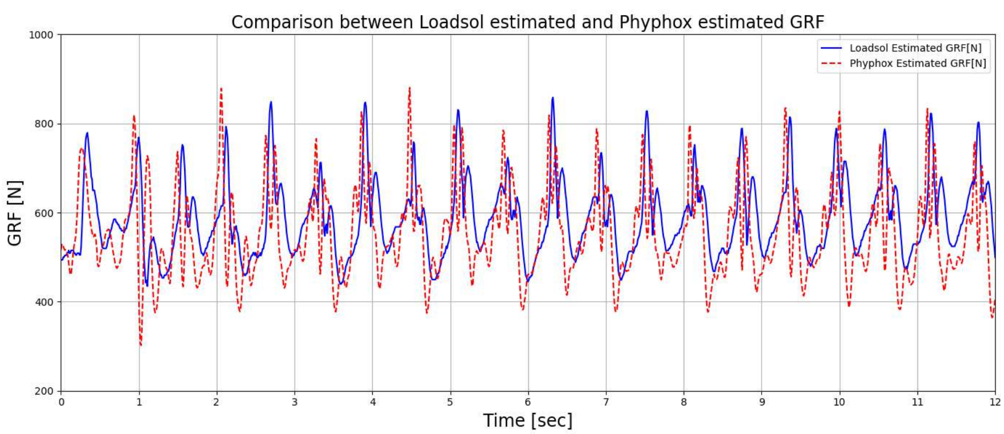

Figure 4.39. Comparison between GRF measured by phone and Loadsol of the first subject.

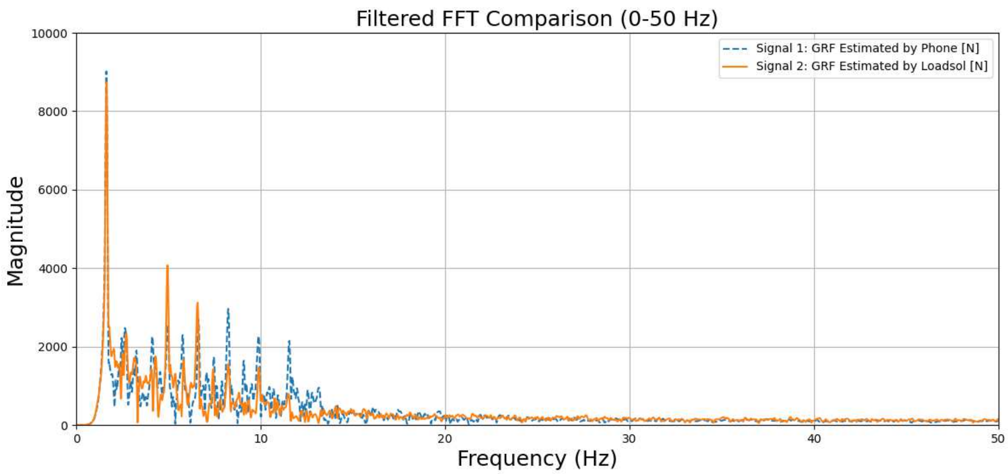

Figure 4.40. The comparison of FFT between Loadsol and Phone collected signals.

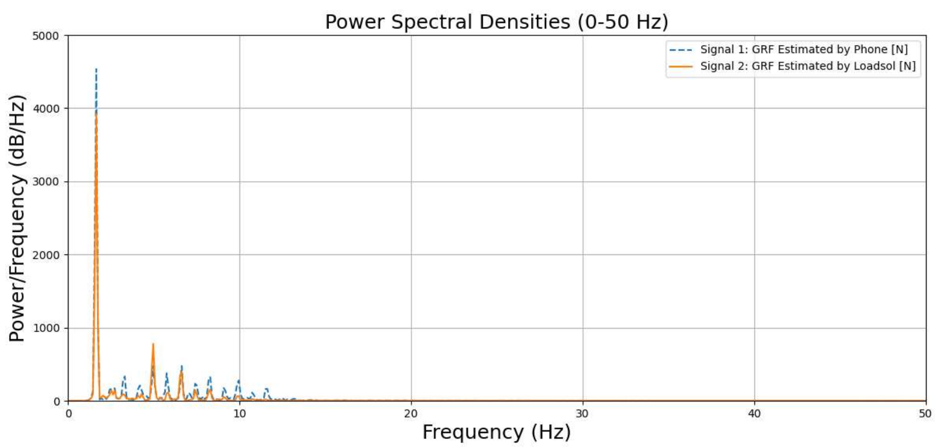

Figure 4.41. The comparison of PSD between Loadsol and Phone collected signals.

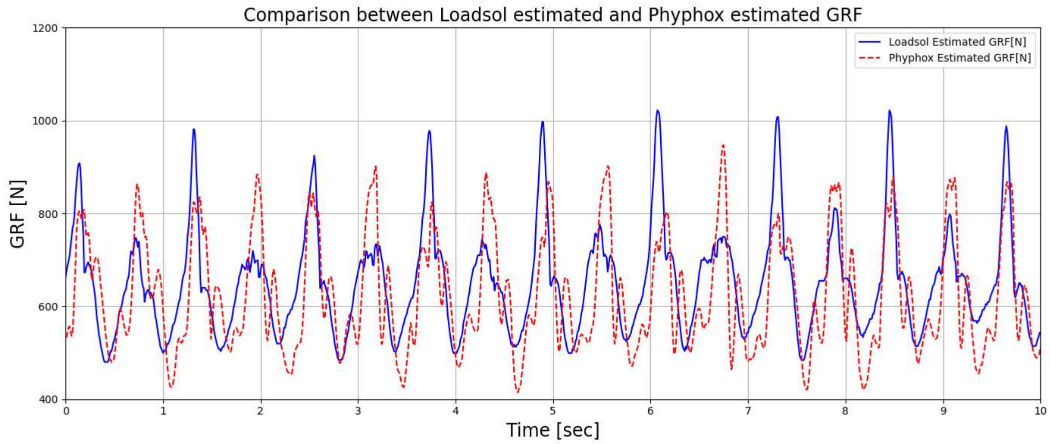

Figure 4.42. Comparison between GRF measured by phone and Loadsol of second subject.

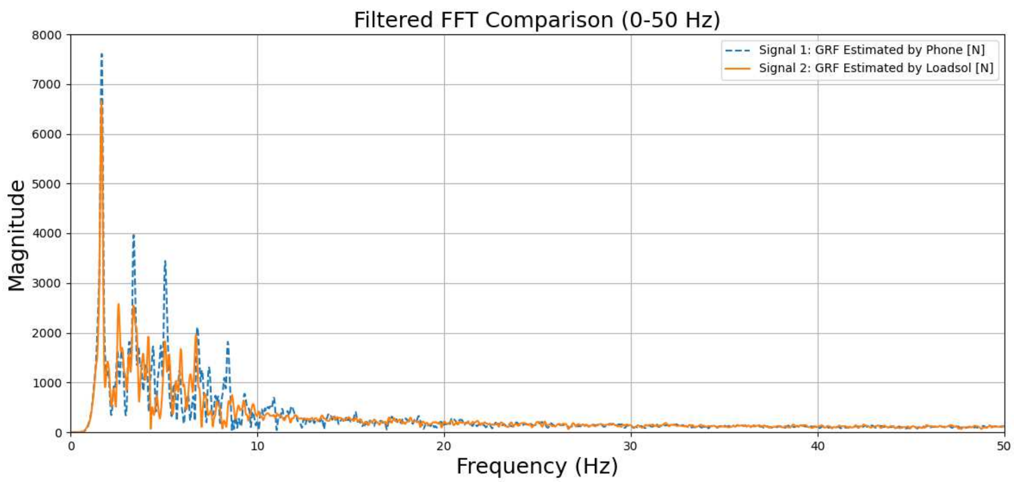

Figure 4.43. Comparison of FFT measured by phone and Loadsol of second subject.

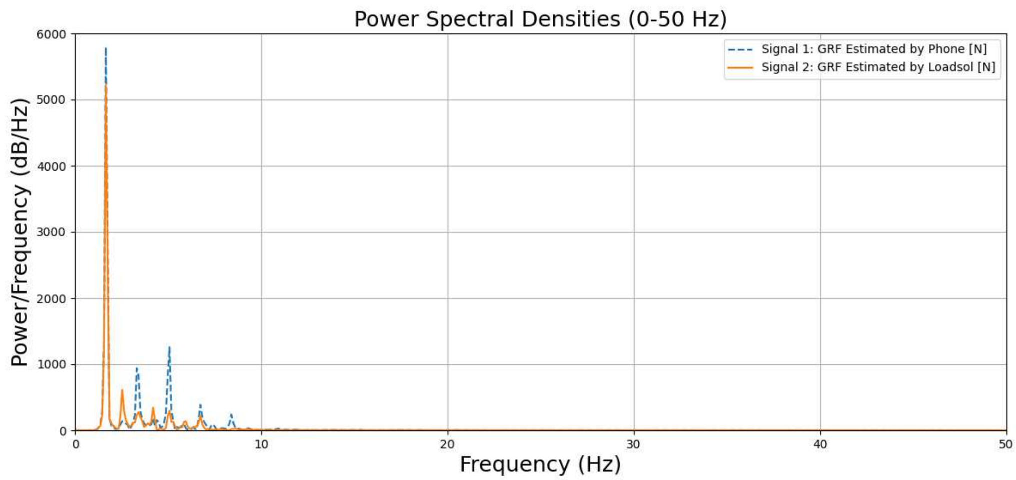

Figure 4.44. Comparison of PSD between two signals of second subject.

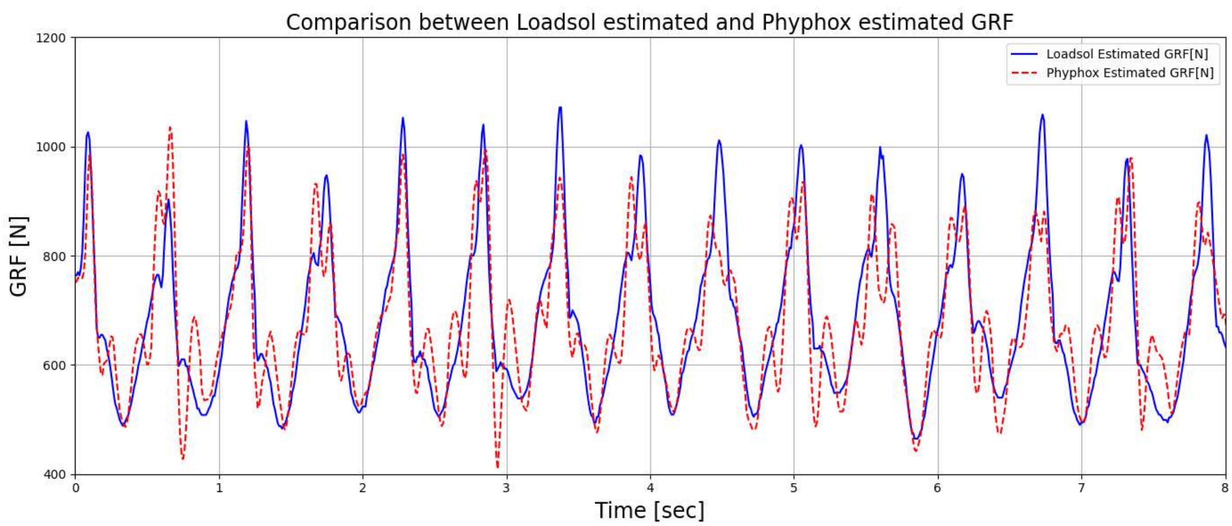

Figure 4.45. Comparison between GRF measured by phone and Loadsol of third subject.

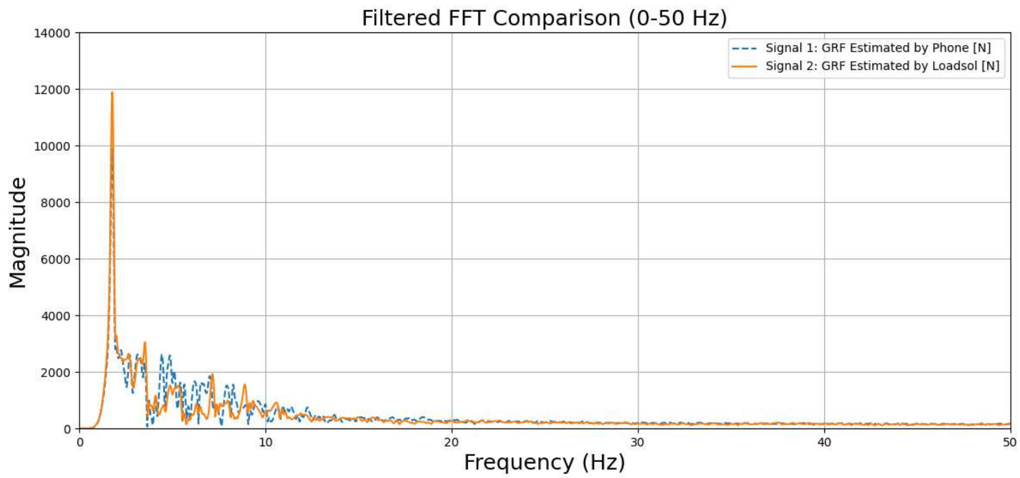

Figure 4.46. Comparison between GRF FFT measured by phone and Loadsol of third subject.

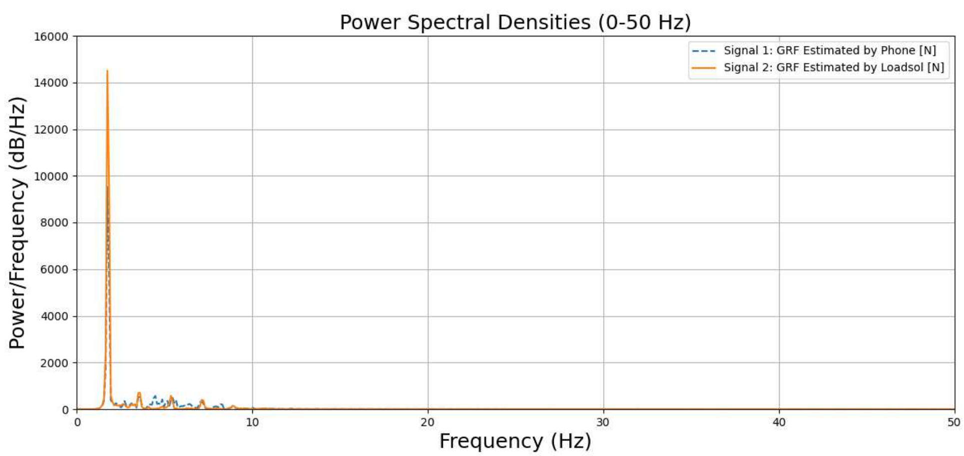

Figure 4.47. Comparison of PSD between two signals of third subject.

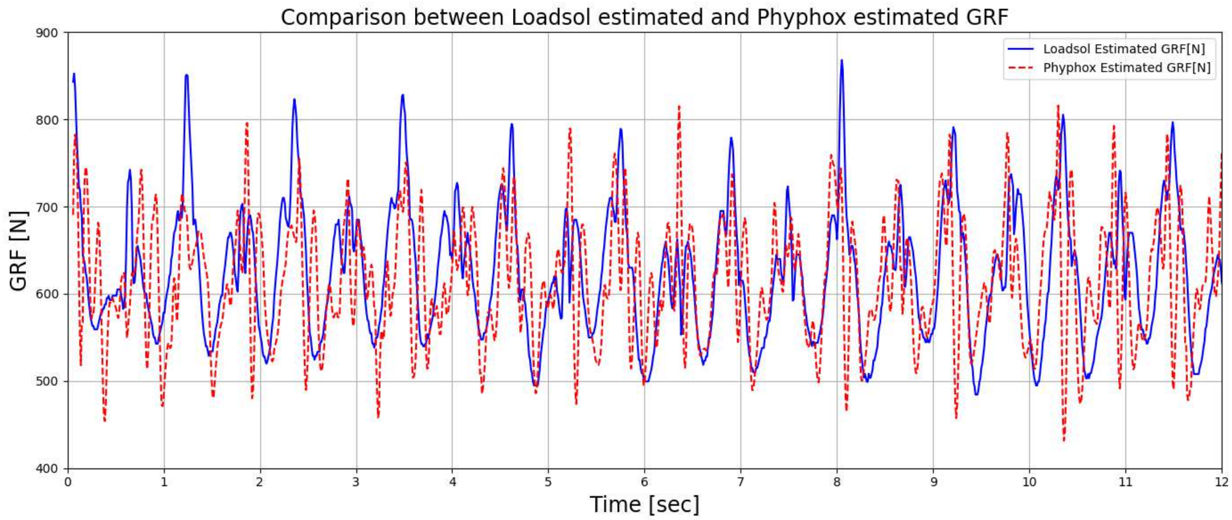

Figure 4.48. Comparison between GRF measured by phone and Loadsol of forth subject.

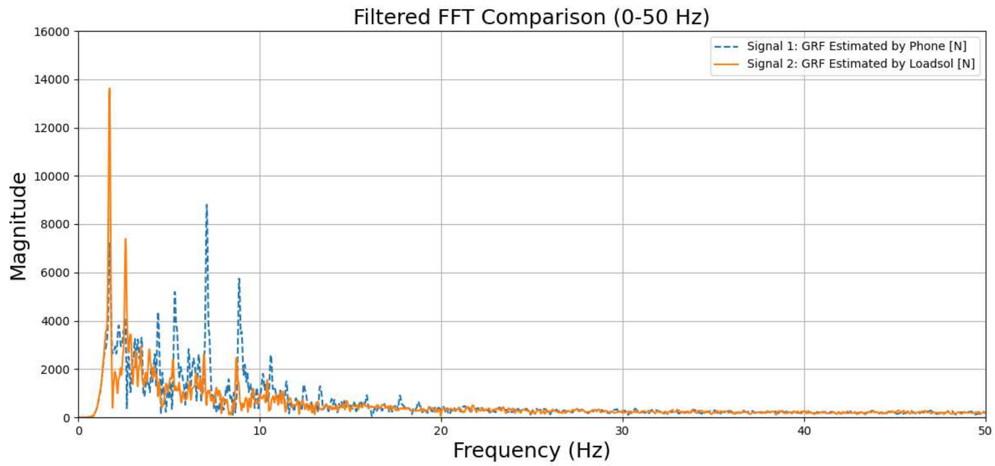

Figure 4.49. Comparison of FFT between two signals of forth subject.

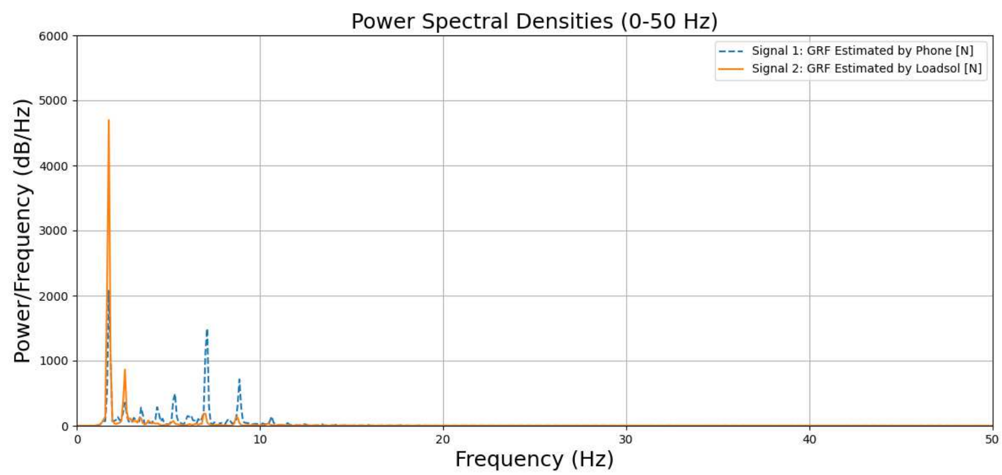

Figure 4.50. Comparison of FFT between two signals of forth subject.

Table 4.4. The objects' alternative impact information for smartphone GRF Estimation.

	Object 2	Object 3	Object 4	Object 5
Height	170 cm	163 cm	168 cm	170 cm
Weight	62 Kg	70 Kg	65 Kg	65 Kg
RMSE	90.7	131.25	76.5	93.03
Standard Deviation of Phone	123.05	235.94	124.46	68.50
Standard Deviation of Loadsol	102.38	218.48	102.27	73.39

As seen from the previous results, the GRF estimation from the phone is still highly correlated with the Loadsol reference at walking frequency, according to PSD and MSC analysis. This observation indicates that mobile devices are accurate enough to measure GRF under various conditions, such as different body weights, center mass, center location, and comfortable walking frequencies.

By comparing several correlation coefficient values, such as RMSE, SD, and Euclidean Distance between smartphone app estimates and Loadsol references, Phyphox shows better performance for GRF estimation in both time and frequency domains. Phyphox also captures GRF peak features well.

In conclusion, mobile device accelerometers are stable and reliable enough to estimate GRF at any frequency on a flat surface. Results collected from mobile devices are highly correlated with the Loadsol reference, especially at walking frequency (comfort frequency). Additionally, there is a significant issue with Loadsol in tracking the first gait load, caused by only the toe or heel touching, or only half of the foot working on the Loadsol, leading to incorrect pressure integration at the first gait. However, mobile devices do not have this issue; they can measure the GRF from the first gait to the last gait perfectly.

4.3.1.4 CV and Smartphone GRF Estimation Metric Parameter Comparison

In Tables 4.3 and 4.4, three parameters: Root Mean Square Error (RMSE) and Standard Deviation (SD) will be analyzed in this chapter.

- The standard deviation of the phone is higher than the standard deviation of the Loadsol reference, indicating a greater spread of the measurements from their respective means. The higher value for the phone means the measurements from the phone are more spread out compared to the data from the Loadsol reference.

- The second parameter discussed in the table is RMSE. This parameter measures the average error between pairs of observations from two signals, indicating how well OpenPose and the smartphone perform compared to the Loadsol reference. Lower RMSE indicates a better correlation between two signals, and vice versa. However, RMSE also depends on the range of values for the two signals. The RMSE value in this research shows how much the OpenPose and smartphone computed value differs from the Loadsol reference and should be considered in relation to the data range. For example, the RMSE for the first object is 90.07, but the overall data range is around 850. Thus, the total error percentage according to the data range is around 10%. Similarly, for the rest of the estimates, the overall tolerance is around 10%.

Overall, the correlation between the Loadsol reference GRF estimation and the two new GRF estimations is not perfectly aligned due to the time series offset and computer vision estimation noise. However, they still exhibit correlation and similarities in GRF estimation.

CHAPTER 5: CONCLUSION

5.1 Conclusion

In this thesis, a CV-based marker-free GRF estimation method has been developed and evaluated by comparing the results with the Loadsol GRF reference. The method's reliability was verified through five objectives. Additionally, a smartphone-based GRF estimation method was thoroughly assessed by collecting walking acceleration and converting it to GRF. The accuracy of the smartphone method was also verified against the Loadsol reference.

In the first part, traditional pressure integration using Loadsol served as the GRF reference, and an CV-based method called OpenPose was involved in this research. Instead of using a depth camera or multiple camera setups, this study employed a single GoPro Hero 9 for key point detection, making it more cost-effective and practical, this method achieved an RMSE of approximately 10%.

In the second part, while the traditional pressure integration using Loadsol served as the GRF reference, smartphones were employed as a new application for acceleration-based GRF estimation. It showed the best correlation with the GRF values computed by the professional Loadsol in the time domain.

In summary, by comparing with previous GRF estimation methods, this research introduces two more cost-effective and high-performance GRF estimation methods: OpenPose and phone accelerometer estimations. The main unique advantages are summarized as follows:

1. Cost-Effectiveness: The two new GRF measurement methods mentioned in previous chapters are more cost-effective than traditional wireless or wearable sensors. The wearable sensors referenced in this research cost around 2000 US dollars. However, the

Phyphox app is free for all users, and the GoPro camera costs only around 200 US dollars, making it more affordable and accessible for researchers.

2. Simplicity and Efficiency: Unlike previous experiments that require multiple cameras and angles, this research uses only one standard camera for GRF estimation. It can effectively monitor displacement in two dimensions, simplifying the process while increasing efficiency.
3. Accurate GRF Analysis: According to the results computed by phone and OpenPose estimation, the real-time GRF can be imported into FEM analysis to predict floor vibration response. This is more accurate than the Fourier 8th polynomial method, which lacks precision and real-time response.
4. Improved Living Comfort: The results can potentially be used for real-time structural response to enhance human living comfort levels, which are often overlooked post-construction. The GRF and floor vibration response can be further analyzed to improve living comfort experiences.

The estimated GRF can be used for human and floor vibration interaction analysis to bridge the gap between the design phase vibration considerations and the actual vibrations occurring in the floor during the post-construction phase. This can enhance human comfort level by addressing real-time walking loads of individual occupants.

5.2 Limitations and Future Work

The first limitation for CV-based GRF estimation is the COCO data detection error. The COCO data-based OpenPose computer vision methods cannot detect human key points 100% accurately at the same location. There is an error of around ±6 millimeters when the video frame rate is 240 frames per second. This high-frequency detection error causes high-frequency noise during computer vision-based GRF estimation. To remove this noise, the signal needs to be processed by several filters, which consequently leads to amplitude loss, identified as the main error in this research. This issue can potentially be resolved by upgrading the OpenPose COCO data or neural networks from OpenCV, which could improve detection accuracy and reduce high-frequency noise.

The second limitation is the time offset caused by the internal clock of the GoPro camera. There were instances of time being ahead or delayed during the test, which resulted in low correlation coefficients. This issue can potentially be solved by applying DTW filters. However, these filters can be aggressive for underestimated signals, leading to a flat period at the end.

The third limitation is that subjects cannot walk on flat ground; instead, they can only walk on the treadmill to avoid displacement errors caused by moving out of the center of the pinhole camera's view, leading to incorrect displacement outputs. Future work can focus on applying a depth camera to monitor key point displacements, allowing subjects to walk on a normal flat ground instead of a treadmill. Additionally, using machine learning can potentially solve this issue. The GRF collected from treadmill walking can serve as a reference, and based on this reference, the signal collected from walking on flat ground can be trained to match the reference, thus outputting similar signals.

The estimated GRF can be imported into simulation systems to analyze the human-floor interaction vibration. Meanwhile, the reference GRF should also be imported into simulation analysis, which can verify the accuracy of smartphone and CV-based GRF (with RMSE around 10% difference with Loadsol reference GRF estimation) for floor response vibration estimation results.

The CV-based GRF estimation is highly recommended in this thesis. In the future, CV-based estimation can work effectively on multiple subjects' GRF estimation by using two non-depth cameras set up on the ground, which is more cost-effective than individual-based smartphone estimation.

Finally, the development of these research methods enables the collection of real-time GRF, which can be imported into simulation system for floor response analysis. Consequently, the floor vibration response can be analyzed to study human living comfort levels in the future.

REFERENCE

Albert, J. A., Owolabi, V., Gebel, A., Brahms, C. M., Granacher, U., & Arnrich, B. (2020). 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. *Sensors*, *20*(18), Article 18. https://doi.org/10.3390/s20185104

Ashburn, M., & Tech, V. (2024). *Architecture professor adapts vibration reduction device for residential use*. EurekAlert! https://www.eurekalert.org/news-releases/1032554

Balageas, D., Fritzen, C.-P., & Güemes, A. (2010). *Structural Health Monitoring*. John Wiley & Sons.

Bobbert, M. F., Schamhardt, H. C., & Nigg, B. M. (1991). Calculation of vertical ground reaction force estimates during running from positional data. *Journal of Biomechanics*, *24*(12), 1095–1105. https://doi.org/10.1016/0021-9290(91)90002-5

Bocca, M., Eriksson, L. M., Mahmood, A., Jäntti, R., & Kullaa, J. (2011). A Synchronized Wireless Sensor Network for Experimental Modal Analysis in Structural Health Monitoring. *Computer-Aided Civil and Infrastructure Engineering*, *26*(7), 483–499. https://doi.org/10.1111/j.1467-8667.2011.00718.x

Bucknall, I., Lane, J., & Palmer, I. (2011). Provisions of the UK National Annex for EN 1991-2 for Rail Bridges. In *Bridge Design to Eurocodes* (pp. 168–182). ICE Publishing. https://doi.org/10.1680/BDTE.41509.0014

Cao, Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(1), 172–186. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2019.2929257

Cao, L., & Chen, J. (2020). Online investigation of vibration serviceability limitations using smartphones. *Measurement*, *162*, 107850. https://doi.org/10.1016/j.measurement.2020.107850

Carroll, S. P., Owen, J. S., & Hussein, M. F. M. (2013). Reproduction of lateral ground reaction forces from visual marker data and analysis of balance response while walking on a laterally oscillating deck. *Engineering Structures*, *49*, 1034–1047. https://doi.org/10.1016/j.engstruct.2012.12.028

Chang, F.-K. (1973). Human Response to Motions in Tall Buildings. *Journal of the Structural Division*, *99*(6), 1259–1272. https://doi.org/10.1061/JSDEAG.0003537

Dallard, P. (2001). *The London Millennium Footbridge*. *79*(22).

Eger, T., Stevenson, J., Boileau, P.-É., & Salmoni, A. (2008). Predictions of health risks associated with the operation of load-haul-dump mining vehicles: Part 1—Analysis of whole-body vibration exposure using ISO 2631-1 and ISO-2631-5 standards. *International Journal of Industrial Ergonomics*, *38*(9), 726–738. https://doi.org/10.1016/j.ergon.2007.08.012

Eguchi, R., Yorozu, A., Fukumoto, T., & Takahashi, M. (2020). Estimation of Vertical Ground Reaction Force Using Low-Cost Insole With Force Plate-Free Learning From Single Leg Stance and Walking. *IEEE Journal of Biomedical and Health Informatics*, *24*(5), 1276–1283. https://doi.org/10.1109/JBHI.2019.2937279

El Kati, R., Forrester, S., & Fleming, P. (2010). Evaluation of pressure insoles during running. *Procedia Engineering*, *2*(2), 3053–3058. https://doi.org/10.1016/j.proeng.2010.04.110

Feldbusch, A., Sadegh-Azar, H., & Agne, P. (2017). Vibration analysis using mobile devices (smartphones or tablets). *Procedia Engineering*, *199*, 2790–2795. https://doi.org/10.1016/j.proeng.2017.09.543

Feng, M., Fukuda, Y., Mizuta, M., & Ozer, E. (2015). Citizen Sensors for SHM: Use of Accelerometer Data from Smartphones. *Sensors*, *15*(2), Article 2. https://doi.org/10.3390/s150202980

Fong, D. T.-P., Chan, Y.-Y., Hong, Y., Yung, P. S.-H., Fung, K.-Y., & Chan, K.-M. (2008). Estimating the complete ground reaction forces with pressure insoles in walking. *Journal of Biomechanics*, *41*(11), 2597–2601. https://doi.org/10.1016/j.jbiomech.2008.05.007

Fortino, G., Guerrieri, A., O'Hare, G. M. P., & Ruzzelli, A. (2012). A flexible building management framework based on wireless sensor and actuator networks. *Journal of Network and Computer Applications*, *35*(6), 1934–1952. https://doi.org/10.1016/j.jnca.2012.07.016

Fuchs, P. X., Hsieh, C.-H., Chen, W.-H., Tang, Y.-S., Fiolo, N. J., & Shiang, T.-Y. (2022). Sensor number in simplified insole layouts and the validity of ground reaction forces during locomotion. *Sports Biomechanics*, *0*(0), 1–14. https://doi.org/10.1080/14763141.2022.2057354

Hajizadeh, M., Clouthier, A. L., Kendall, M., & Graham, R. B. (2023). Predicting vertical and shear ground reaction forces during walking and jogging using wearable plantar pressure insoles. *Gait & Posture*, *104*, 90–96. https://doi.org/10.1016/j.gaitpost.2023.06.006

Hanagan. (1997). Active control approach for reducing floor vibrations. *Journal of Structural Engineering*, *123*(11), 1497–1505. https://doi.org/10.1061/(ASCE)0733-9445(1997)123:11(1497)

Helbing, D., Buzna, L., Johansson, A., & Werner, T. (2005). Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions. *Transportation Science*, *39*(1), 1–24. https://doi.org/10.1287/trsc.1040.0108

Jin, L., Zhang, J., Hold-Geoffroy, Y., Wang, O., Blackburn-Matzen, K., Sticha, M., & Fouhey, D. F. (2023). Perspective Fields for Single Image Camera Calibration. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 17307–17316. https://doi.org/10.1109/CVPR52729.2023.01660

Jones, A. J., & Saunders, D. J. (1972). Equal comfort contours for whole body vertical, pulsed sinusoidal vibration. *Journal of Sound and Vibration*, *23*(1), 1–14. https://doi.org/10.1016/0022-460X(72)90785-7

Ju, M., Park, C., & Kim, G. (2015). Structural Health Monitoring (SHM) for a cable stayed bridge under typhoon. *KSCE Journal of Civil Engineering*, *19*(4), 1058–1068. https://doi.org/10.1007/s12205-015-0039-3

Kang, J.-D., Baek, E.-R., & Park, S.-H. (2023). Evaluation of Vibration Detection Using Smartphones in a Two-Story Masonry-Infilled RC Frame Building. *Buildings*, *13*(4), Article 4. https://doi.org/10.3390/buildings13041069

Kim, Gu, Na, Park, Kim, & Kim. (2020). Wireless Ground Reaction Force Sensing System Using a Mechanically Decoupled Two-Dimensional Force Sensor. *IEEE/ASME Transactions on Mechatronics*, *25*(1), 66–75. IEEE/ASME Transactions on Mechatronics. https://doi.org/10.1109/TMECH.2019.2948650

Kim, Pakzad, S., Culler, D., Demmel, J., Fenves, G., Glaser, S., & Turon, M. (2007). Health monitoring of civil infrastructures using wireless sensor networks. *Proceedings of the 6th International Conference on Information Processing in Sensor Networks - IPSN '07*, 254. https://doi.org/10.1145/1236360.1236395