Reconstruction of the respiratory rate using wrist-based PPG, accelerometer and gyroscope

Master of Science (Tech) Thesis
University of Turku
Department of Computing
Health Technology
June 2025
Piita Koivisto

Supervisors:
Kianoosh Kazemi
Antti Airola

UNIVERSITY OF TURKU
Department of Computing

Piita Koivisto: Reconstruction of the respiratory rate using wrist-based PPG, accelerometer and gyroscope

Master of Science (Tech) Thesis, 54 p.
Health Technology
June 2025

Breathing is an important indicator of serious illnesses. Traditional measurement methods are often uncomfortable and not suitable for everyday use. Consequently, there is a growing need for accurate and easily adoptable solutions for continuous respiratory rate monitoring. Wearable approaches that utilize biosignals have emerged as a promising way to address this need. However, biosignals often contain a significant amount of noise, which can reduce the accuracy of the measurements.

A wide range of methods have been proposed for respiratory rate estimation, including both traditional signal processing techniques and modern machine learning approaches. In this thesis, we utilize accelerometer, gyroscope, and photoplethysmography (PPG) signals collected from the wrist to develop a pipeline for accurate respiratory rate estimation. The goal is to produce a pipeline that combines waveform analysis techniques, such as the Fast Fourier Transform and modulation extraction methods, to enable efficient estimation without relying solely on computationally expensive machine learning algorithms. However, we do employ machine learning methods to assess signal quality as accurately as possible, since signal quality has a significant impact on subsequent analysis. An additional objective is to assess whether the fusion of estimates from multiple modalities yields improved accuracy relative to the use of PPG alone. We adopted a pipeline proposed in previous research as a foundation and explored ways to improve its performance.

The pipeline based on PPG produced results comparable to those of other methods reported in the literature; however, it did not outperform the most recent deep learning approaches. Although the fusion method achieved similar accuracy, it was more computationally intensive. Therefore, relying solely on the PPG pipeline appears to be a more efficient and accurate approach. Notably, the fusion pipeline performed better than the baseline pipeline adopted from previous work.

Keywords: PPG, accelerometer, gyroscope, respiration, wearables

Contents

Abbreviations

RIP	Respiratory inductance plethysmography
PPG	Photoplethysmograpy
DC	Constant component
RIIV	Respiratory-Induced Intensity Variation
RIAV	Respiratory-Induced Amplitude Variation
RIFV	Respiratory-Induced Frequency Variation
ACC	Accelerometer
GYR	Gyroscope
BrPM	Breaths-per-minute
PCA	Principal Component Analysis
ICA	Independent Component Analysis
FFT	Fast Fourier Transform
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
RR	Respiratory rate
IRR	Instantaneous respiratory rate
SST	Synchrosqueezing transform
AR	Autoregressive model
MAE	Mean absolute error
OCSVM	One-class support vector machine
PC1	First principal component
MA	Motion activity
WLS	Weighted least squares
RMSE	Root mean square error

1 Introduction

Breathing is one of the most important physiological functions in humans. It is vital for gas exchange, which removes carbon dioxide produced by the body and allows for the intake of fresh oxygen. Ventilation also supports important social functions such as laughter and speech production. [1], [2] Respiratory rate (RR) can be used to assess patient wellness and to identify various medical conditions. Abnormal respiration may indicate health deterioration and can be a sign of cardiac issues or respiratory conditions such as pneumonia. [3]

Many traditional RR measurement methods, such as respiratory inductance plethysmography (RIP) belts, have been shown to be both inaccurate and uncomfortable for use in everyday settings. As a result, there is increasing interest in the use of consumer-grade wearable devices to support more accessible and unobtrusive health tracking. Recent research [4], [5] has highlighted smartwatches as promising platforms for RR monitoring. These devices are easy to use, relatively affordable, and already integrated into the daily routines of millions of people. Modern smartwatches are equipped with multiple sensors, such as accelerometers (ACC), gyroscopes (GYR) and photoplethysmography (PPG), which can be utilized for measuring respiratory activity. Unlike specialized medical devices, smartwatches offer convenience and social acceptability, which may improve adoption and continued use over time. Importantly, smartwatches operate on flexible, programmable platforms that support third-party applications. This opens up the possibility of

deploying respiratory monitoring capabilities through software alone, without requiring changes to the hardware. For many users, such features may operate in the background alongside other health-related functions, including heart rate tracking, sleep monitoring, and activity recognition. This versatility contributes to the potential value of smartwatches as practical tools for health monitoring in everyday, non-clinical settings. [6]

The objective of this thesis is to develop and test an algorithm that leverages PPG, ACC and GYR signals measured from the wrist to provide accurate RR estimation. We focus on waveform analysis techniques because they are highly efficient, offering low runtime and reduced energy consumption with minimal CPU usage compared to machine learning approaches [7]. We define waveform analysis as a set of signal processing methods that extract physiological information directly from the shape, amplitude, frequency, and timing characteristics of the signals, without relying on computationally intensive machine learning models. We adopted the pipeline proposed by Huang *et al.* [8] and later evaluated by Kazemi *et al.* [9] as a starting point, and explored modifications to improve its accuracy. We also explored if we could simplify the pipeline to improve its computational efficiency. We utilize techniques such as time-frequency decomposition (e.g., Fourier and wavelet transforms), peak detection, and interval analysis to identify patterns associated with respiration-induced variations in the signals. This approach is particularly suitable for smartwatches and other wearable devices, where limited processing power and battery capacity make lightweight, real-time methods essential for continuous physiological monitoring.

1.1 Research questions

The objective of this thesis is to explore the following research questions:

RQ1: Can an algorithm that integrates PPG and motion-based data such as ACC and GYR produce accurate RR estimates?

RQ2: Can this be done by only using waveform analysis methods?

RQ3: Does incorporating motion-based RR estimates enhance the accuracy of RR estimation derived from PPG signals?

2 Background

Breathing is regulated by the autonomic nervous system. The process consists of two main phases: inhalation, during which oxygen is drawn into the lungs and diffuses across the alveolar membrane into the bloodstream; and exhalation, during which carbon dioxide, resulting from cellular metabolism, is expelled from the body. [10] Respiration is a permanent and dynamic biosignal which means that it does not need an artificial trigger to be measured but it can change due to physiological or environmental reasons. [11] Research has demonstrated that a change in RR can serve as an early warning sign for identifying high-risk patients, often up to 24 hours before a critical event occurs. [12] Compared to blood pressure or heart rate, RR has been shown to be a more reliable predictor of potential clinical deterioration. [13]

2.1 Classic ways to measure RR

The simplest way to measure RR is by manual observation. This approach is often used in emergency situations when a quick estimate is needed. However, it can be challenging—especially if the person's breathing is shallow or irregular. The results may also vary depending on the method of observation. For example, listening with a stethoscope can produce a different outcome than visually counting chest movements. Additionally, if the observation period is short and the person is aware of being monitored, they may unintentionally change their breathing to appear calmer

or deeper. [14]

In research settings, respiration is typically measured from the airways. This involves placing a mask—or a nose clip and mouthpiece—around the nose and mouth to form an airtight seal and ensure accurate measurements. [12] However, studies have shown that wearing a mask can alter natural breathing patterns. People tend to take deeper, slower breaths when using a mask, which increases tidal volume and reduces respiratory frequency. [15] RIP belt is commonly used as an indirect method for tracking RR and chest volume changes. While effective, modern RIP systems often require up to an hour to stabilize before providing accurate data, and they may be uncomfortable for long-term wear. [16] These limitations highlight the need for a more practical, non-invasive, and reliable approach to continuous respiration monitoring.

2.2 Biosignals in respiration detection

There are several different biosignals that can be used to measure RR. In this thesis, we focus on PPG, ACC, and GYR signals.

2.2.1 PPG

PPG is a non-invasive method that uses light to detect changes in blood volume beneath the skin. [17] These variations are caused by a pressure pulse that the heart creates as it pumps blood towards the peripheral circulation. This final pulse intensity is minuscule compared to the pulse that the heart creates but it can be measured.

The PPG signal is typically measured using a pulse oximeter, composed of a light source and a photodiode receiver. There are two main methods for measuring PPG: transmission and reflectance. In transmission PPG, the receiver is positioned

directly across the tissue from the light source and measures how much light passes through the tissue—for example, measurements taken from the fingertip. In reflectance PPG, the receiver is positioned next to the light source and detects the light reflected back from the tissue, as is common in wrist-based measurements. The amount of the transmitted or reflected light is analyzed to detect changes in blood volume within the tissue. [18] Beer-Lambert law states that an absorbing medium reduces light intensity exponentially, and this chance can be detected by the receiver photodiode. Absorbance is heavily influenced by the optical density of the measuring site, the distance the light travels, and the blood volume present in the tissue at any given time. This variation in blood volume over time can be captured using PPG. [19] The most common light sources are infrared or green wavelength. Green is the most suitable option for wrist-based sensors, as it has been observed to penetrate deeper into the skin than other wavelengths, which has led to it providing the most accurate readings.

The most common measurement sites are the fingertip, forehead and earlobe. However, these locations require additional devices to be attached to the person. A wrist-worn PPG device would be the most convenient way to measure the signal, as it is comfortable and easy for the patient to use. It would be even more convenient if the measurement could be performed using a device the patient already wears. Commonly used smartwatches fit this purpose well. [17]

Respiration variation in PPG

There are two components in the PPG signal: a pulsatile component and a constant (DC) component. The pulsatile component exhibits a waveform that reflects fluctuations in blood volume, appearing as peaks and troughs in the signal. This component is primarily used to determine heart rate. In contrast, the DC appears as modulations in the baseline of the signal. Variations in these components can be

used to estimate RR.

Respiration induces several variations in the PPG signal, each affecting different aspects of the waveform. One of these variations is *Respiratory-Induced Intensity Variation (RIIV)*, which influences the DC component of the PPG. This effect arises due to pressure changes in the thoracic cavity, which alter venous return—the rate of blood flow back to the heart. During inhalation, intrathoracic pressure decreases, leading to a reduction in central venous pressure and an increase in venous return. During exhalation, the opposite occurs, resulting in baseline modulations in the PPG signal. Another important variation is *Respiratory-Induced Amplitude Variation (RIAV)*, which affects the pulsatile component of the PPG signal. Intrathoracic pressure changes influence stroke volume, reducing the quantity of blood the left ventricle ejects during inspiration. Consequently, the PPG pulse amplitude decreases during inhalation and increases during exhalation, creating a rhythmic fluctuation in amplitude. In addition to these variations, respiration also induces *Respiratory-Induced Frequency Variation (RIFV)*, which is regulated by the autonomic nervous system. This variation occurs due to respiratory sinus arrhythmia, a well-documented phenomenon where heart rate increases during inhalation and decreases during exhalation. This natural modulation of heart rate is often used as an indirect marker of respiratory activity. [18]

Each of these variations provides valuable insight into respiration, but as Li *et al.* [20] concluded, no single modulation is consistently reliable for accurate respiratory analysis. Instead, the most precise estimation of RR is achieved by combining multiple modulation estimates, as this approach accounts for different physiological influences and improves robustness in varying conditions. The modulations are illustrated in Fig. 2.1.

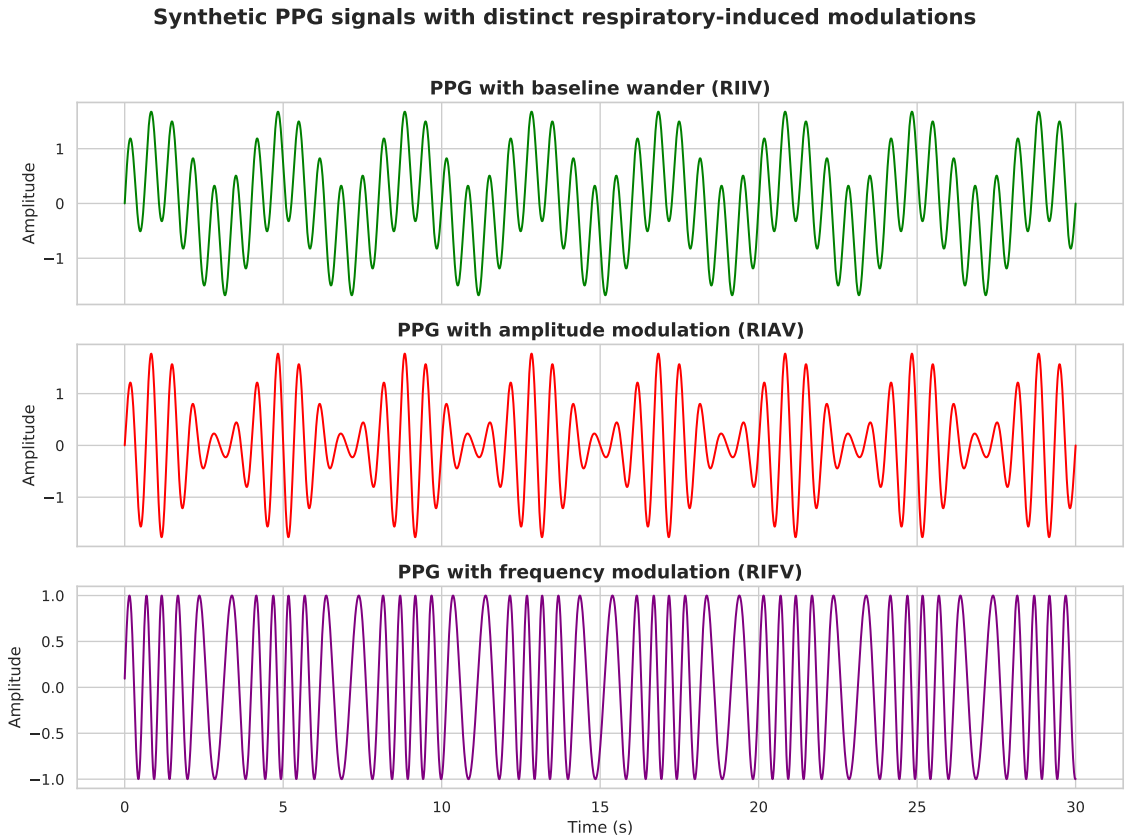

Figure 2.1: Synthetic PPG signals illustrating the three primary respiration-induced modulations. Top: Baseline wander caused by thoracic pressure changes modulating the DC component. Middle: Amplitude variation due to respiratory effects on stroke volume. Bottom: Frequency variation reflecting respiratory sinus arrhythmia. Each signal shows an isolated and amplified modulation effect for enhanced visual clarity.

PPG challenges

One of the issues with PPG signals is its inability to provide accurate readings during the movement of daily life. This interference is called motion artifacts which usually comes from hand movement when utilizing a wrist worn monitor. [17] Other source of noise in PPG is usually high frequency noise originating from instrumentation. The effectiveness of PPG measurement is also impacted by skin tone and how securely the device is attached to the skin. Higher melanin levels in darker

skin absorb more light, thereby reducing the intensity of the detected signal. If the PPG device is not in secure contact with the skin at the measuring site, ambient light interference can affect the signal quality. [8], [21] Persons age, sex and different disorders can affect the PPG signal. Some diseases such as diabetes or sleep apnea can affect autonomic activity, which in turn affects RIFV. It is also observed that different sexes have different variations in the PPG. The day-to-day health of the person can also have an effect on the signal. For example, dehydration or loss of blood volume due to blood donation can cause measurements to differ and it can be seen both in RIIV and RIAV. [22]

2.2.2 Motion sensors

ACC and GYR are common motion sensors used in wearable devices to track body movements. These sensors are compact, energy-efficient, and capable of detecting the small physical movements associated with breathing, making them useful for estimating RR in everyday settings. [9]

ACC measures proper acceleration—how quickly the velocity of an object changes relative to free fall. This is different from coordinate acceleration, which is measured in a fixed external frame. ACC captures motion along three axes (x, y, and z), and the combined signal can be used to estimate overall movement while minimizing the effect of the sensor's orientation [23]. In smartwatches and fitness trackers, ACC can detect the slight movements at the wrist caused by chest and abdominal motion during breathing. However, these signals are often weak and can be affected by other movements, so measurements are typically more accurate when the person is at rest. [5]

GYR measures angular velocity, or how fast something is rotating, also along three axes. It is often included in an inertial measurement unit, along with ACC, to give a fuller picture of motion [24]. For respiration monitoring, a wrist-worn

GYR can pick up small, repeated rotational movements linked to breathing. These movements are caused by the natural motion of the upper body during inhalation and exhalation and travel through the arm to the wrist, where they can be recorded. GYR signals could be less sensitive to linear movement noise than ACC, making them a useful complementary tool for estimating RR [5].

2.3 Signal processing

Raw signal data collected from these devices contain information related to various biological processes, along with external interference such as noise. The raw data must first be processed to extract the relevant modulation patterns associated with respiratory activity.

2.3.1 Signal segmentation

Biosignals are typically recorded over extended periods, making direct analysis challenging. To facilitate more manageable and reliable processing, the signal is divided into smaller, fixed-duration segments or windows. This segmentation assumes that both the RR and noise characteristics remain relatively stable within each window. The segment length should be long enough to capture the meaningful features of the signal, but short enough to maintain approximate stationarity [9]. Research has suggested that segment lengths between 30 and 90 seconds are suitable for estimating RR [25]. Longer segments improve the accuracy of the estimates [19], while shorter ones make the computations more efficient and stable.

2.3.2 Filtering

Filtering is a crucial preprocessing stage in signal processing, used to eliminate noise and extract relevant frequency components. One widely used approach is But-

terworth filtering, chosen for its favourable frequency response characteristics. A Butterworth filter is engineered to provide a smooth frequency response across the passband, free of the ripple characteristic of Chebyshev and elliptic filters. The filter was originally introduced in 1930 by Stephen Butterworth in his seminal paper "On the Theory of Filter Amplifiers" [26]. A bandpass Butterworth filter allows frequencies within a specific range to pass while attenuating frequencies outside that range. This makes it useful in isolating physiological signals of interest. Respiration signals typically occupy a frequency range of approximately 0.2 to 0.6 Hz, corresponding to breathing rates of 12 to 36 breaths-per-minute (BrPM). This range excludes both low-frequency components associated with baroreflex variation (typically 0.1–0.15 Hz), which reflects short-term blood pressure regulation, and higher-frequency components related to cardiac activity, which generally occur above 0.8 Hz. [27]

2.3.3 Quality estimation

Quality control is essential for identifying and filtering out noisy or corrupted portions of the signal that could compromise the accuracy of downstream analyses. Signal quality can be evaluated using a range of methods—statistical, spectral, or machine learning-based—depending on the characteristics of the sensor data. These methods typically rely on features such as signal variability, kurtosis, shape consistency, or measures of rhythmic structure, like peak intervals or peak-to-valley distances. Based on these features, threshold-based rules can be used to classify segments as either acceptable or unreliable [7], [8]. In more advanced approaches, machine learning models are trained to recognize abnormal signal patterns without requiring explicit thresholds [28]. Instead of discarding entire signal segments, some techniques apply quality control at a finer resolution using short, overlapping subwindows. This allows for the retention of clean portions of data while excluding only the parts affected by brief artifacts [7]. These approaches have been applied to PPG

and motion signals, and have been shown to improve the stability and accuracy of respiration estimates.

2.3.4 Dimensionality reduction techniques

Motion sensors such as ACC and GYR typically measure signals along three orthogonal axes—x, y, and z. In many wearable applications, it is useful to reduce these three signals to a single representative signal that captures the dominant motion component. This simplification can improve the robustness of subsequent analysis by reducing the influence of sensor orientation and noise from irrelevant directions. A common and straightforward approach is to compute the total magnitude of motion from the triaxial signal. For an ACC, this is calculated as:

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{2.1}$$

where a_x, a_y, and a_z are the acceleration values along each axis. This method produces a single, positive-valued signal representing the overall intensity of motion, regardless of direction. [23] While simple and effective, this approach may combine signals from different sources, including noise.

To better isolate the signal component related to respiration, Principal Component Analysis (PCA) can be applied [8]. PCA is a statistical technique that identifies the directions (principal components) in which the data varies the most. By projecting the original 3D signal onto the first principal component (PC1)—which captures the highest variance—a one-dimensional signal can be extracted that emphasizes the dominant motion pattern, which is assumed to correspond to respiration. PCA is particularly useful when the sensor orientation is unknown or varies over time. Another method for dimension reduction is Independent Component Analysis (ICA), which assumes that the observed signals are mixtures of statistically independent

source signals. In the context of motion sensing, ICA can help separate respiration-related movement from other motion sources, such as arm gestures. By decomposing the multi-axis signal into independent components, ICA may identify a component that closely corresponds to the respiration signal. This component can then be selected as the one-dimensional representation for further analysis. [9] While both PCA and ICA can be effective for extracting respiration-related features, their performance depends on the characteristics of the input signals.

2.4 Respiratory wave extraction methods

Several signal processing techniques have been applied to extract respiratory waveforms from physiological signals. These include traditional frequency-domain approaches such as Fourier and wavelet analysis, as well as other methods like autoregressive modelling [29] and data-driven machine learning techniques [9]. In this section, we focus on two commonly used methods: the Fast Fourier Transform (FFT) and the Continuous Wavelet Transform (CWT).

2.4.1 Fourier analysis

The FFT is a cornerstone algorithm in signal processing, used to convert signals from the time domain into the frequency domain. While physiological signals like respiration are typically recorded as time-varying waveforms, analyzing their frequency components can help reveal patterns related to their periodic behaviour. The FFT is an efficient algorithm for computing the *Discrete Fourier Transform (DFT)*, which mathematically decomposes a sequence of N sampled data points $x[n]$ into a sum of sinusoidal components at different frequencies. The DFT is defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N}, \quad \text{for } k = 0, 1, \ldots, N-1 \tag{2.2}$$

where, $X[k]$ represents the amplitude and phase of the sinusoidal component at the k^{th} frequency bin, $x[n]$ is the input signal in the time domain, j is the imaginary unit, and N is the total number of samples.

Computing the DFT directly requires $O(N^2)$ operations, which is computationally intensive for large data sets or real-time analysis. The FFT, introduced by Cooley and Tukey in 1965 [30], significantly improves this by reducing the complexity to $O(N \log N)$ through the use of mathematical symmetries and redundancies in the DFT computation. This increased efficiency makes the FFT especially valuable in practical applications such as the continuous monitoring of RR using wearable sensors. By applying the FFT to ACC data, the dominant frequency component—corresponding to the cyclic motion of breathing—can be extracted and used to estimate the RR (see Fig. 2.2).

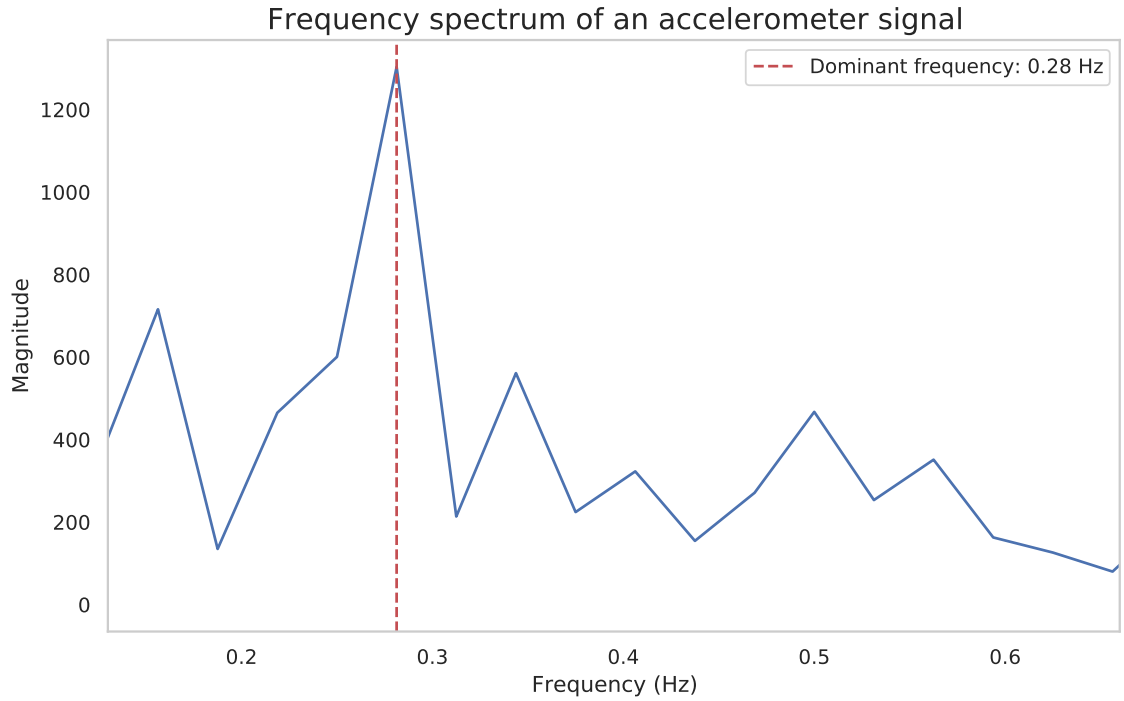

Figure 2.2: Frequency spectrum of a 32-second window of ACC signal, computed using the FFT. The red line indicates the dominant frequency component, corresponding to the estimated RR of 16.8 BrPM.

2.4.2 Wavelet analysis

Wavelet transform has become a widely adopted tool in biomedical signal processing due to its effectiveness in analysing non-stationary signals, such as PPG. Traditional techniques like the FFT assume stationarity and provide only global frequency information, making them less reliable for dynamic physiological signals. In contrast, wavelet transform preserves both time and frequency information, making it ideal for detecting transient features and localized frequency changes. Wavelet Transform works by decomposing a signal into a series of scaled and shifted versions of a function denoted as the mother wavelet. These transformations allow the signal to be represented in the frequency domain while retaining its time-domain characteristics. The wavelet coefficients, obtained through this decomposition, reflect how the signal correlates with the wavelet at different scales and positions. By linearly combining these weighted wavelet functions, the original signal can be accurately reconstructed. Furthermore, wavelet transform is capable of de-noising and compressing signals without introducing significant distortion.

Among the various wavelet techniques, the CWT and the Discrete Wavelet Transform (DWT) are most commonly used. DWT is more computationally efficient and is often preferred for real-time applications or where noise reduction is a priority. [31] However, DWT provides only discrete scale and time representations, which may limit its time-frequency resolution. The CWT, by contrast, performs a continuous convolution of the input signal with scaled and shifted versions of the mother wavelet. This produces a high-resolution time-frequency map, known as a scalogram, which shows how different frequency components of a signal evolve over time. The CWT is defined as:

$$\mathrm{CWT}_x(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\, \psi * \frac{(t-b)}{a}\, dt \tag{2.3}$$

where a is the scale, b is the time shift, and $\psi(t)$ is the mother wavelet. Although CWT is computationally more intensive, its advantages in resolution and interpretability make it particularly well-suited for analyzing physiological signals where precision and clarity are the priority. [32]

3 Related work

In recent years, considerable research has focused on improving the accuracy and convenience of RR monitoring. A growing body of work explores noninvasive techniques, particularly those employing motion-based or light-based sensors. Among these approaches, wrist-worn devices, such as smartwatches, have gained popularity due to their widespread availability, user comfort, and ease of integration into daily life [5]. Several studies have investigated the use of motion-based sensors, such as ACC and GYR, to detect respiration-related movements [4], [5], [23], [33]. However, other researchers have demonstrated that light-based methods, particularly those utilizing PPG, can yield more accurate estimations of RR [19], [22]. In this chapter, we review a collection of studies conducted between 2006 and 2023 that utilize PPG or motion sensors in combination with waveform analysis techniques for RR estimation.

3.1 Datasets

Most datasets used for RR estimation have been collected under controlled laboratory conditions [5], [7], [8], [19]. These datasets typically involve a small number of healthy participants, ranging from 10 to 15 individuals [5], [23], [34]. A few studies have employed moderately larger datasets. For example, Dai *et al.* [7] collected data from 32 participants using wrist-based PPG during various activities, while Huang *et al.* [8] included 38 participants, recording PPG and ACC data from the upper

arm during sitting and supine conditions. In contrast, Leube *et al.* [33] collected a significantly larger dataset of wrist-worn ACC recordings from 223 individuals, including patients with sleep-related disorders, in a clinical sleep laboratory.

In addition to newly collected datasets, some researchers have leveraged publicly available resources. Dehkordi *et al.* [22] utilized the CapnoBase dataset, which contains two subsets: a test set of forty-two 8-minute recordings from 42 patients (29 children and 13 adults) undergoing general anesthesia, and a calibration set of 124 two-minute segments used to fine-tune algorithm parameters. Each recording contains synchronized ECG, PPG, and capnometry signals. Capnography is used as the reference standard for RR, with breaths manually annotated by a research assistant. Several studies have also used the MIMIC II database, which contains data from patients in the intensive care unit. For example, Pimentel *et al.* [29], [35] selected a subset of 53 adult patients from Beth Israel Deaconess Medical Center in Boston. The dataset includes PPG and respiratory signals recorded via impedance plethysmography, with analysis focused on 8-minute segments of stable, spontaneous breathing.

Most datasets used for RR estimation are relatively small and collected under controlled conditions, often from healthy subjects. While studies have begun to include larger or more diverse populations—particularly through clinical datasets or public repositories—there is a need for broader, real-world datasets to improve the generalizability of RR estimation methods. A real-world dataset collected during an instrument validation study [36] was employed by Kazemi *et al.* [9] for RR estimation and will likewise be used in this thesis.

3.2 Experiment setups

There are multiple approaches to measuring RR. Motion-based methods, however, typically require the subject to be in a resting state to ensure reliable measurements

[34]. Several studies [5], [34] have evaluated different body positions, such as lying down, sitting, or standing, to determine the most suitable posture for accurate measurement. Hernandez *et al.* [5] found that sitting and supine positions yielded more accurate RR estimates compared to standing, particularly when using a wrist-worn GYR.

Similarly, PPG-based respiration estimates are often affected by noise, most commonly in the form of motion artifacts. In [7], researchers developed RespWatch, a system designed to provide robust RR estimation from wrist-worn PPG sensors, even in the presence of noise and motion. To evaluate its effectiveness under realistic conditions, they collected data from 32 participants performing a range of tasks intended to elicit varying degrees of cognitive load, emotional stress, and physical movement. These tasks included watching a video, preparing and delivering a speech, solving mathematical problems, and undergoing a cold pressor test. The study demonstrated that the type of task impacted the RR estimation accuracy. Notably, periods involving increased movement, such as math solving and free time periods, yielded larger errors, while low-motion tasks such as watching a video, resulted in more accurate RR estimates. This highlights the importance of adaptive estimation methods for wearable RR monitoring in everyday life.

3.3 Signals and devices

In the literature, RR estimation approaches can generally be grouped into those that use a single sensor and those that combine multiple sensors. Single-sensor methods are preferred for their simplicity, low power requirements, and ease of integration, particularly in wearable applications. While efficient, these systems may be more susceptible to noise or signal degradation, especially in less controlled environments. To mitigate these challenges, some studies have explored multi-sensor approaches that combine information from different sources. This section outlines both single-

and multi-sensor methods for RR estimation, with a focus on their signal processing techniques, sensor fusion strategies, and reported performance in the literature.

3.3.1 Single sensor

To maintain simplicity in analysis, some studies have relied on a single sensor. For example, Haescher *et al.* [23] employed a wrist-worn ACC to estimate the RR. The sensor recorded acceleration along three axes (x, y, and z), which were then combined into a single signal using the magnitude of the acceleration vectors. The study highlighted several advantages of a single-sensor setup, including low power consumption and simplified sensor placement. Additionally, a single ACC was suggested to be more energy efficient than PPG, as its power consumption is significantly lower (ACC: 0.5–2 mW; PPG: 1–50 mW).

Leube *et al.* [33] extracted additional features from the ACC data beyond the x, y, and z axes. Specifically, they calculated two orientational angles: (1) the angle between the y-axis and the x-z plane and (2) the angle between the x-axis and the gravity acceleration vector. These features were projected to serve as even more reliable respiration proxies than the traditional acceleration axes. The underlying rationale was that wrist movements detected by the ACC are not directly caused by respiration itself but rather by the subtle mechanical rotation of the wrists, which occurs as the body moves naturally during inhalation and exhalation. This rotation could also be detected using GYR, which we will be utilizing in this work. The most common way to extract RR from motion based signals is to utilize FFT. [5], [23], [34]. Signals are first normalized and filtered to the wanted frequency band. The filtered signals are then transformed to the frequency domain using FFT. The respiration frequency is determined to be the highest amplitude frequency in the respiration band.

PPG has been widely employed for estimating RR as a stand-alone modality

[19], [22], [29]. When used independently, PPG enables non-invasive RR monitoring by analyzing respiratory-induced modulations in the signal. Dehkordi *et al.* [22] proposed a method to extract the instantaneous respiratory rate (IRR) from the PPG signal through a three-stage process: (1) the synchrosqueezing transform (SST), a time-frequency analysis technique that combines wavelet analysis with a method that sharpens the time-frequency representation by moving signal components to more accurate locations, was used to isolate the respiratory-induced variations from the PPG signal; (2) SST was applied again to each of these extracted components to estimate their respective IRRs; and (3) a peak-conditioned fusion technique was employed to combine the individual estimates into a final IRR measurement. Similarly, Karlen *et al.* [19] developed a real-time RR estimation method for PPG. Their algorithm extracted the three respiratory-induced variations using the Incremental-Merge Segmentation algorithm [37]. The dominant frequencies of these components were identified using FFT. To enhance reliability, they proposed a Smart Fusion approach that combined the individual frequency estimates using a transparent mean, selectively discarding unreliable measurements based on artifact detection and inter-estimate disagreement.

Pimentel *et al.* [29] improved the robustness of RR estimation by applying multiple autoregressive (AR) models to the three respiratory-induced variations. In this method, each variation is modeled using an "all-pole" AR filter, where the filter's poles correspond to resonant frequencies. The dominant respiratory frequency is determined by identifying the pole with the largest magnitude whose phase angle falls within the physiologically plausible respiratory frequency range. This approach enhances accuracy by leveraging model-based spectral estimation, particularly under noisy or artifact-prone conditions. Pimentel *et al.* [35] used an Gaussian process with a periodic covariance function to extract the RR from the PPG variations. This method provides a confidence measure which is used to combine the estimates

into a single RR estimate. Confidence measure is employed in the weighted average algorithm which uses bigger weights for RR estimates with smaller variances. The method also improves accuracy by using only valid windows of the signal by determining a signal quality index. The index is calculated by utilizing template matching and flat-line detection to determine which windows are unsuitable for reliable RR estimations.

3.3.2 Multiple sensors

Several studies have explored the use of multiple sensors to improve the accuracy of RR estimation [5], [8]. By combining signals from different modalities, these methods aim to capture complementary physiological information and improve the accuracy of the methods.

ACC and GYR

Hernandez *et al.* [5] evaluated various sensor combinations to estimate the RR using ACC and GYR data collected from the wrist in a laboratory setting. First, to reduce noise, each signal component (x, y or z axis from either GYR or ACC) was smoothed using an averaging filter. The component with the highest periodicity, defined by the peak amplitude within the 0.13–0.66 Hz frequency range, was selected as the representative respiratory waveform. RR was then estimated in the frequency domain by identifying the dominant peak using the FFT. To integrate estimates from multiple sensors, a weighted average was computed based on the normalized magnitudes of their respective dominant frequency components.

Their results showed that all sensor types, GYR, ACC, and their combination, achieved good accuracy, with mean errors below one breath per minute. Notably, the GYR alone provided the most accurate measurements (mean error of 0.38 BrPM), outperforming the ACC (mean error of 0.97 BrPM). The authors attributed this to

wrist rotation caused by chest movement during breathing, which the GYR captures more effectively. Interestingly, fusing both sensors did not enhance accuracy (mean error of 0.55 BrPM), likely due to the fusion algorithm's reliance on selecting the most periodic signal component, which may not consistently benefit from signal combination. Furthermore, the GYR consistently outperformed the ACC across all body positions, reinforcing this finding. These results highlight the importance of carefully designing fusion strategies. Improper signal integration may reduce accuracy instead of improving it.

PPG and ACC

In addition to combining motion sensors, other studies have proposed hybrid approaches that integrate physiological and motion-based signals. For example, Huang *et al.* [8] proposed a method for estimating RR using both PPG and ACC data collected from the upper arm in a laboratory setting. Their algorithm consisted of two parallel processing pipelines. First, respiration-induced modulations were extracted from the PPG signal, and RR was computed as the reciprocal of the mean duration of the 30 most recent inner breath intervals during low-motion periods. In the second pipeline, the three orthogonal ACC signals were projected onto their PC1 to capture the dominant motion pattern. FFT and peak tracking were then applied to estimate RR in the frequency domain. The final RR estimate was obtained by combining the PPG- and ACC-based results, with signal quality used to weight each contribution. If both modalities yielded low-quality estimates, the final RR value was discarded to prevent unreliable output. The method demonstrated reasonable performance, with a mean absolute error of 2.3 $\pm$ 1.4 BrPM, a bias of –0.8 BrPM, and an outage rate of 30%. This hybrid approach demonstrates how integrating physiological and inertial signals can enhance robustness, provided that signal quality assessment is incorporated into the fusion strategy.

Kazemi *et al.* [9] tested this approach using the real-world dataset from [36], with the difference that they used ICA instead of PCA to combine the ACC axis. They also included GYR signals, in addition to PPG and ACC signals, in the pipeline to improve accuracy. Their approach yielded slightly lower performance with MAE of 2.90 ± 0.70 BrPM and mean bias of -2.2 BrPM. This may be due to the inclusion of additional GYR signals and the fact that Huang *et al.* [8] collected their dataset under controlled laboratory conditions from the upper arm, which is less prone to noise compared to the wrist-based data collection used by Kazemi *et al.* [9].

3.4 Summary of related work

Recent research on RR estimation has shifted from traditional airflow and chest movement measurements to non-invasive methods using wearable sensors. Wrist-worn devices, leveraging motion sensors like ACC and GYR, and light-based sensors such as PPG, have become increasingly popular for their convenience and accessibility. Motion-based methods work well in low-activity settings but are sensitive to noise, while PPG-based techniques can offer higher accuracy by detecting respiratory-induced signal variations. Several studies have proposed hybrid systems that combine motion and physiological signals to improve robustness, often incorporating signal quality assessments to adapt to real-world conditions. Publicly available datasets like CapnoBase and MIMIC II, as well as custom-collected lab datasets, are commonly used to evaluate these approaches. Overall, the literature highlights the trade-offs between sensor simplicity, power consumption, and estimation accuracy, emphasizing the importance of careful sensor fusion and noise handling.

4 Data

This thesis utilizes a dataset collected as part of an instrument validation study [36] to investigate whether wrist-based signals recorded during everyday activities can be used to estimate RR. The dataset comprises PPG, ACC, and GYR signals recorded from the non-dominant wrist of each participant. To evaluate the accuracy of wrist-based RR estimation, a chest-worn ACC served as a reference signal.

4.1 Participants and Recruitment

Participants were recruited in southern Finland between July and August 2019 using convenience sampling and a snowball recruitment approach. To ensure a homogeneous sample, exclusion criteria included restrictions on wearable device usage during work, physical activity limitations, cardiovascular disease, and the presence of illness symptoms at the time of data collection.

4.2 Ethics

Ethical considerations were addressed in accordance with the Declaration of Helsinki and the Finnish Medical Research Act (No 488/1999), and approval was obtained from the University of Turku for Human Sciences Ethics Committee. Participants received both oral and written information about the study before providing informed consent. Participation was voluntary, and individuals could withdraw at

any time without providing a reason.

4.3 Data collection

The study involved 45 healthy individuals aged 18–55 years, with an equal distribution of sexes. Participants wore the study devices continuously for seven days while engaging in their normal daily activities. For this analysis, we focused on data from the first 24 hours, as participants also wore a Shimmer3 ACC on their chest during this time to provide reference data for evaluating our methods. An overview of the data collection setup is provided in Fig. 4.1.

The Shimmer3 is a lightweight, noise-resilient device designed for continuous triaxial ACC data recording. It features sufficient internal storage and battery life to operate for 24 hours without interruption. The device was set to record at a sampling rate of 512 Hz, capturing detailed motion data for reference. A chest-worn ACC is a validated method for accurately detecting respiration [38]. Wrist-based data were collected using the Samsung Gear Sport smartwatch, which recorded PPG, triaxial GYR, and ACC signals at a sampling frequency of 20 Hz. The smartwatch is a light and waterproof device with a battery life of up to three days, which makes it suitable for everyday use.

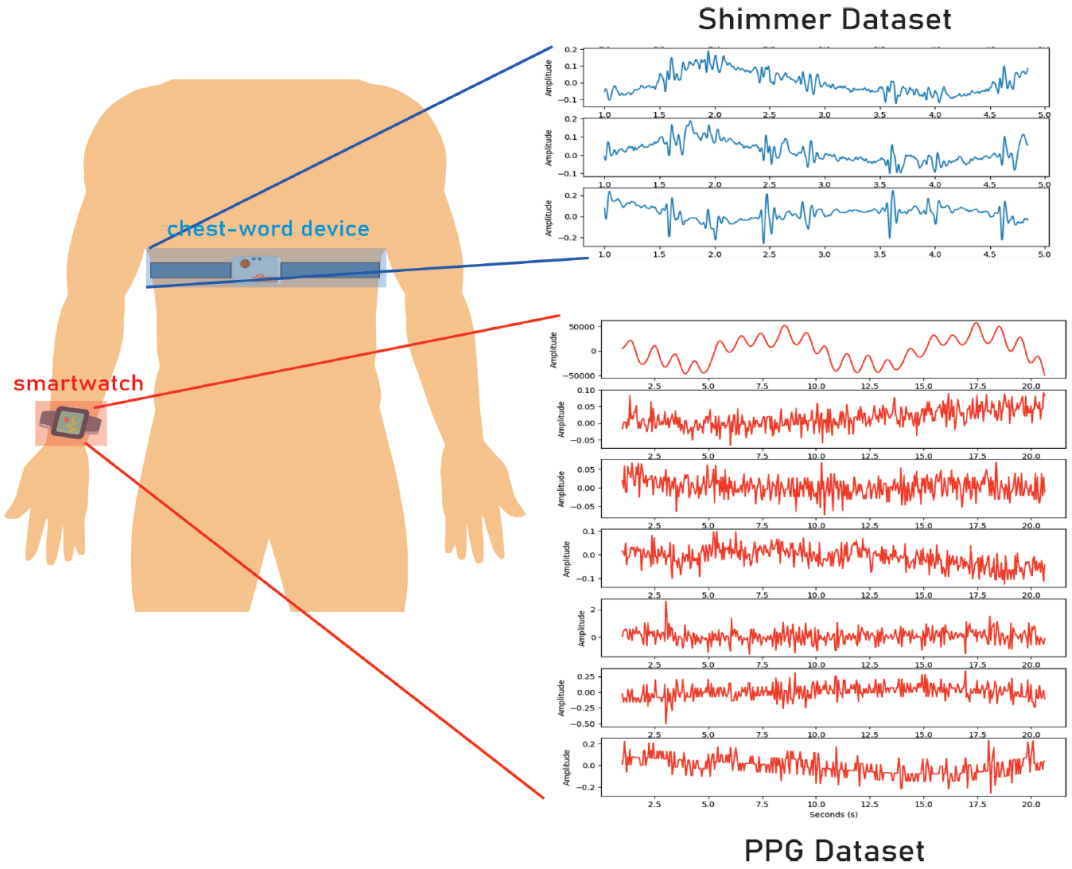

Figure 4.1: Data collection setup featuring a Samsung Gear Sport smartwatch and a chest-mounted Shimmer3 tri-axial ACC. The smartwatch recorded PPG, tri-axial ACC, and tri-axial GYR signals, while the Shimmer3 device provided reference motion data from the chest. [9]

5 Methods

We develop and test a signal processing pipeline for estimating RR from ACC, GYR, and PPG signals. The accuracy of the pipeline is evaluated against ground-truth RR values obtained from reference data. We adopted the pipeline proposed by Huang *et al.* [8] and later evaluated by Kazemi *et al.* [9] as a starting point, and explored modifications to improve its accuracy. This pipeline was selected because it utilizes waveform analysis techniques and has been tested on the same dataset used in this thesis.

5.1 Preprocessing

We followed the preprocessing steps outlined by Kazemi *et al.* [9]. To ensure reliable RR estimation, all signals were segmented into non-overlapping 32-second windows prior to analysis. To enhance the resolution of wrist-based PPG, ACC, and GYR signals, they were upsampled from 20 Hz to 64 Hz using linear interpolation. The method estimates new data points by fitting a straight line between existing values and placing additional points at regular intervals based on the upsampling rate. This increases the signal's temporal resolution while preserving its continuity. [39] For ACC and GYR the upsampling was done separately for each of the three axis. We also decided to filter the ACC and GYR axis using a Butterworth bandpass filter with cut-off values of 0.13 and 0.66 Hz as in [5] due to difficulties in subsequent analysis. PPG signals were filtered similarly with a Butterworth bandpass filter

with cut-off values of 0.1 and 1 Hz.

5.2 Reference data

The triaxial ACC data collected from the chest, which served as the reference in this analysis, was processed in three main steps as described by Kazemi *et al.* [9]. First, for each 32-second window, the ACC signals from the x, y, and z axes were preprocessed to reduce noise and enhance respiration-related modulations. Then an FFT was applied separately to each axis to analyze the frequency content of the signals. The RR for each axis was then estimated by identifying the dominant frequency peak below 0.5 Hz, following the removal of the DC component. Finally, the three individual estimates were fused using a Kalman filter to obtain the final RR.

5.3 PPG

The RR was extracted from the PPG signal through a two-step process: (1) signal quality assessment and (2) RR estimation. To explore potential improvements to the pipeline proposed by Huang *et al.* [8], we employed two distinct methods for RR extraction: the modulation-based approach used in their work, and the CWT.

5.3.1 Quality control

PPG signals are often contaminated by noise, which can affect the accuracy of subsequent analysis. To ensure signal reliability, we performed a signal quality assessment prior to further processing using a model developed and trained by Feli *et al.* [28]. Kazemi *et al.* [9] used this approach when evaluating the pipeline, and we also adopted it due to its demonstrated precision, low execution time, and minimal energy consumption, making it suitable for low-power devices such as the smartwatch

used in this study. The model is a one-class support vector machine (OCSVM) that evaluates the quality of each signal window based on five features, classifying them as either reliable or unreliable.

Quality assessment features

Feli *et al.* [28] described three of the five features as both highly accurate and computationally efficient: (1) The range of energy across heart cycles quantifies inter-cycle energy variability. Consistent values suggest stable, high-quality signals, while large variations may indicate the presence of noise or artifacts. The energy of a heart cycle is defined as the sum of the squared amplitudes of the signal over the duration of a single cardiac cycle:

$$E_{\text{hc}} = \sum_{n=1}^{N} |x_{\text{hc}}(n)|^2 \tag{5.1}$$

where $x_{\text{hc}}(n)$ represents the amplitude of the PPG signal at the n-th sample of a single heart cycle, and N is the total number of samples in that heart cycle.

(2) The average Euclidean distance between a template and individual heart cycles captures morphological similarity by measuring how closely each heartbeat matches the shape of a template cycle. The template is defined as the averaged cardiac cycles of the signal. A low Euclidean distance implies that the beat closely resembles the expected template waveform, indicating signal reliability. The average Euclidean distance is computed as:

$$D_{\text{avg}} = \frac{1}{M} \sum_{i=1}^{M} \sqrt{\sum_{j=1}^{l} \left(hc_{i,j} - \text{template}_j \right)^2} \tag{5.2}$$

where D_{avg} is the average Euclidean distance between the template and M heart cycles in a segment, each of length l. $hc_{i,j}$ is the j-th sample of the i-th heart cycle, and template_j is the corresponding sample in the template.

(3) The average correlation between the template and individual heart cycles provides another measure of morphological consistency, where higher correlation values indicate more regular and coherent waveform shapes. Each heart cycle is compared to the template using the Pearson correlation coefficient, and the results are then averaged across all cycles within the segment, as shown in Eq. 5.3.

$$r_{\text{avg}} = \frac{1}{M} \sum_{i=1}^{M} \frac{\sum_{j=1}^{l} \left(hc_{i,j} - \bar{hc}_i\right) \left(\text{template}_j - \bar{\text{template}}\right)}{\sqrt{\left(\sum_{j=1}^{l} \left(hc_{i,j} - \bar{hc}_i\right)^2\right) \left(\sum_{j=1}^{l} \left(\text{template}_j - \bar{\text{template}}\right)^2\right)}} \tag{5.3}$$

where r_{avg} is the average Pearson correlation coefficient across all M heart cycles in the segment. The term $hc_{i,j}$ denotes the j-th sample of the i-th heart cycle, and $\bar{hc}_i$ is its mean value. The reference template is denoted by template_j, which represents the j-th sample of the template waveform, and $\bar{\text{template}}$, its mean.

Additionally, Feli *et al.* [28] employed two supplementary features with low computational cost: the interquartile range, which measures variability in the central portion of the signal and flags segments with high variability as unreliable; and the standard deviation of the power spectral density, which identifies segments with broad frequency content as unreliable.

OCSVM

Signal segments are classified as either reliable or unreliable using a pre-trained OCSVM. This model uses the previously described features to detect segments that deviate from the characteristics of clean, high-quality PPG signal. Reliable signals tend to be consistent and show low variation, while unreliable signals are typically more dispersed and irregular. The OCSVM had been trained using only high-quality (reliable) data. Specifically, the model mapped the training data into a high-dimensional feature space using a kernel function and learned a boundary that

encloses the reliable training samples with the maximum possible margin. Importantly, the model was trained only on reliable data, making it well-suited for situations where the specific characteristics of unreliable signals are unknown. During testing, both reliable and unreliable segments were evaluated: test samples falling within the learned boundary were classified as reliable, while those outside it were considered unreliable. This method allows for the detection of noisy or anomalous signals without needing explicit labels for all types of artifacts.

We used this trained model to evaluate signal quality by calculating the proportion of reliable segments within each 32-second analysis window. A window was considered to be of acceptable quality if at least 90% of its segments were classified as reliable. Windows falling below this threshold were excluded from further analysis. The difference between high- and low-quality signals can be seen in Fig. 5.1.

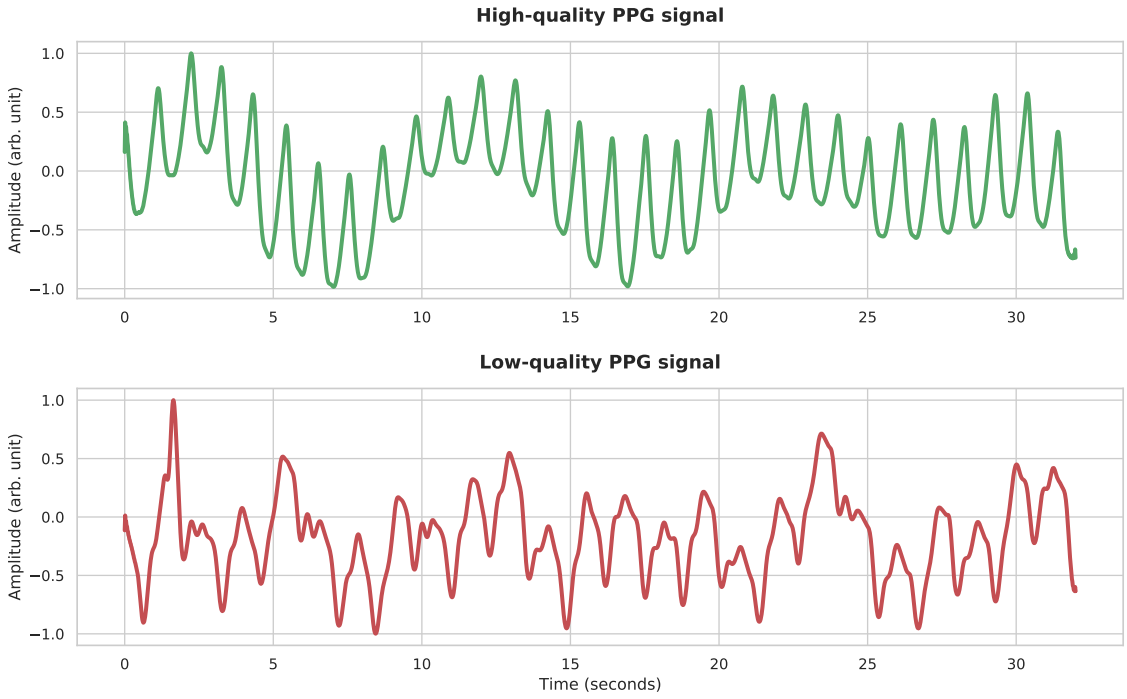

Figure 5.1: Examples of high- and low-quality PPG signals.

5.3.2 Modulation based RR estimation

To evaluate the performance of the original pipeline, we estimated the RR using the modulation-based method described by Huang *et al.* [8]. The algorithm consists of three main steps: (1) extraction of systolic peaks from the PPG signal, (2) derivation of the three respiration-induced modulations based on these peaks, and (3) calculation of the final RR estimate as the average of the three individual estimates.

Systolic Peak Detection

The PPG signal was first filtered using a 3th-order Butterworth bandpass filter with a passband of 0.5–8 Hz to isolate cardiac-related components. We then applied a systolic peak detection method based on the approach proposed by Elgendi *et al.* [40]. First the filtered signal was rectified by zeroing negative values and then squared to enhance the prominence of systolic peaks. Next, two moving averages were calculated to help identify regions of interest: a short-term average (111 ms) corresponding to the typical duration of a systolic pulse and a long-term average (667 ms) to approximate the baseline. A dynamic threshold was calculated by offsetting the long-term average by a fraction of the mean signal energy. Segments where the short-term average exceeded this threshold were considered candidate pulse waves. Within each segment, the local maximum was identified, and the most prominent peak was selected as the systolic peak. To avoid false detections of closely spaced peaks within a single cardiac cycle, a minimum inter-peak interval of 300 ms was enforced between successive detections.

Extraction of respiration-induced modulations

Three respiration-induced modulations were extracted using the systolic peak information.

Respiratory-Induced Frequency Variation (RIFV): The timestamps of the detected systolic peaks were used to compute inter-beat intervals. Let t_i denote the timestamp of the i-th systolic peak. The inter-beat interval (IBI) sequence is defined as:

$$\mathrm{IBI}_i = t_{i+1} - t_i \tag{5.4}$$

These intervals were interpolated onto a uniform time axis using cubic spline interpolation to produce a smooth continuous RIFV signal. A bandpass filter (0.1–1 Hz) was then applied to isolate the respiratory-frequency components.

Respiratory-Induced Amplitude Variation (RIAV): Using the same peak timestamps, the corresponding peak amplitudes were extracted. A smooth amplitude variation curve was obtained using cubic spline interpolation and subsequently filtered in the 0.1–1 Hz band, similar to RIFV.

Respiratory-Induced Intensity Variation (RIIV): RIIV curve, representing slow baseline shifts in the PPG signal due to respiration, was extracted by applying a low-pass Butterworth filter with a cutoff frequency of 0.15 Hz directly to the PPG signal.

After extracting the three modulated signals seen in Fig. 5.2, RR estimation was performed on each by detecting the positive and negative peaks within each curve. A thresholding method was used to eliminate insignificant or overly similar peaks. The number of valid breaths within a window was then determined, and the average breath duration was computed by dividing the interval between the first and last detected peaks by the total number of breaths. This duration was converted into BrPM to obtain an RR estimate for each modulation type.

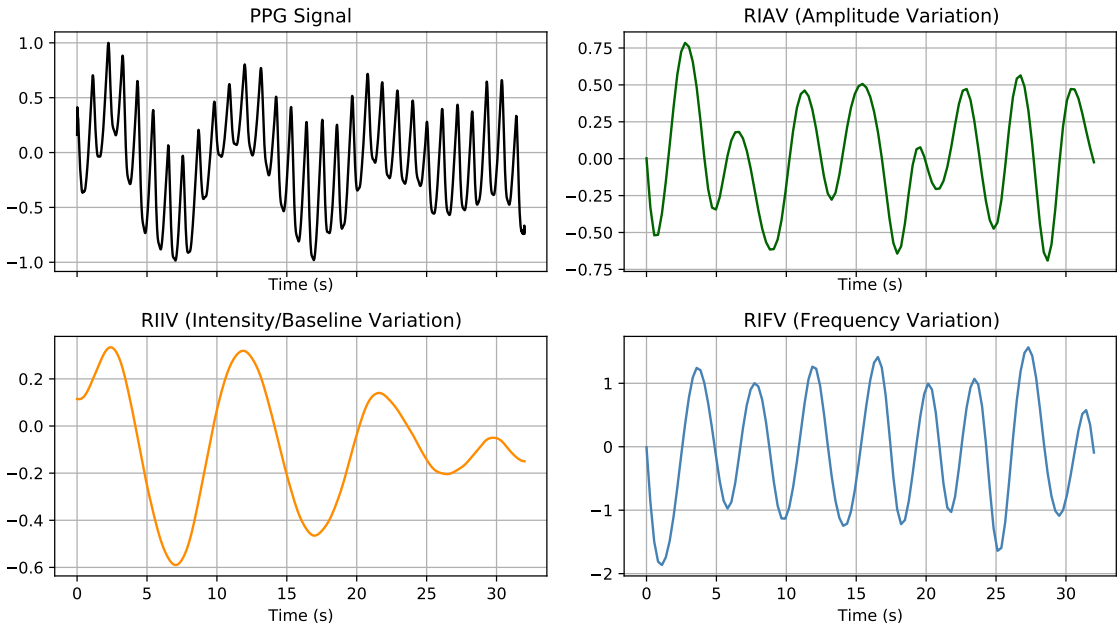

Figure 5.2: PPG signal and the respiration-induced modulations.

Final RR Estimate

The final RR estimate for each time window was obtained by averaging the three individual estimates (RIFV, RIAV, RIIV).

5.3.3 CWT

To see if we can improve the accuracy of the pipeline, we tested a CWT method based on the Morlet wavelet to estimate the RR from PPG. The algorithm begins by defining a frequency band of 0.13 to 0.7 Hz, which corresponds to typical human breathing rates. These frequencies are converted into appropriate wavelet scales for the CWT. The Morlet wavelet is then used to compute the CWT of the PPG signal, yielding a matrix of complex coefficients across time and frequency. To quantify the energy content at each time-frequency point, the magnitude of the complex coefficients is squared, producing a scalogram as seen in Fig. 5.3. The total power for each frequency is then computed by summing across the time dimension.

The frequency with the maximum total power, the strongest consistent presence throughout the window, is identified as the dominant respiratory frequency.

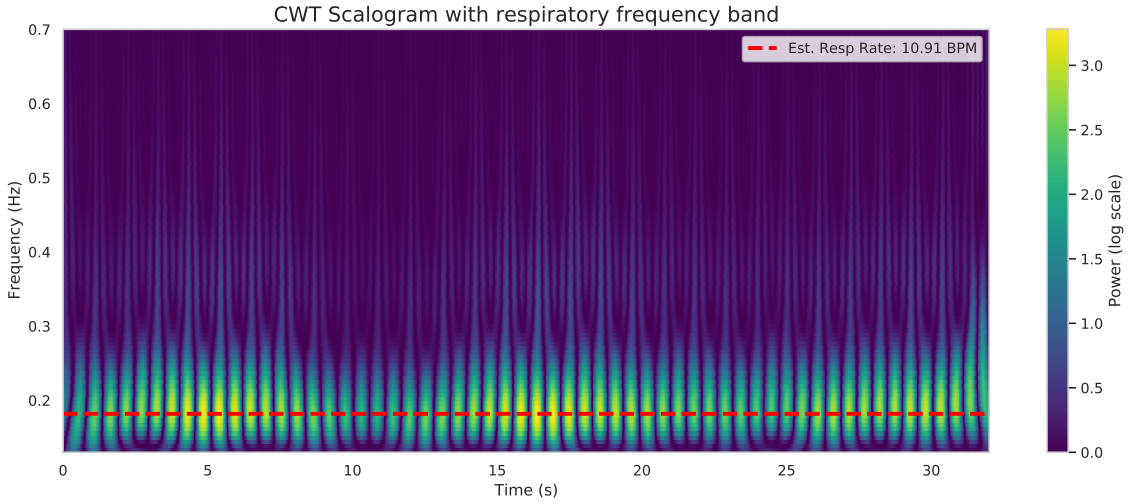

Figure 5.3: Time–frequency representation of the PPG signal using CWT. The scalogram displays the wavelet power spectrum (log-scaled) over time, highlighting respiratory-frequency activity. The red dashed line indicates the estimated dominant respiration frequency, corresponding to a rate of approximately 10.91 BrPM.

5.4 ACC and GYR

The RR was extracted from the ACC and GYR signals through a three-step process: (1) axis combination, (2) signal quality estimation, and (3) RR estimation.

5.4.1 Combining the axis

To detect respiration-induced movement in the ACC and GYR signals, the x, y, and z axes of each window were combined using PCA, following the approach by Huang *et al.* [8]. We also evaluated ICA method proposed by Kazemi *et al.* [9] for this purpose, but it resulted in lower performance; therefore, PCA was selected.

The respiratory component was assumed to be captured primarily in the PC1. This reduction in dimensionality can be seen in Fig. 5.4.

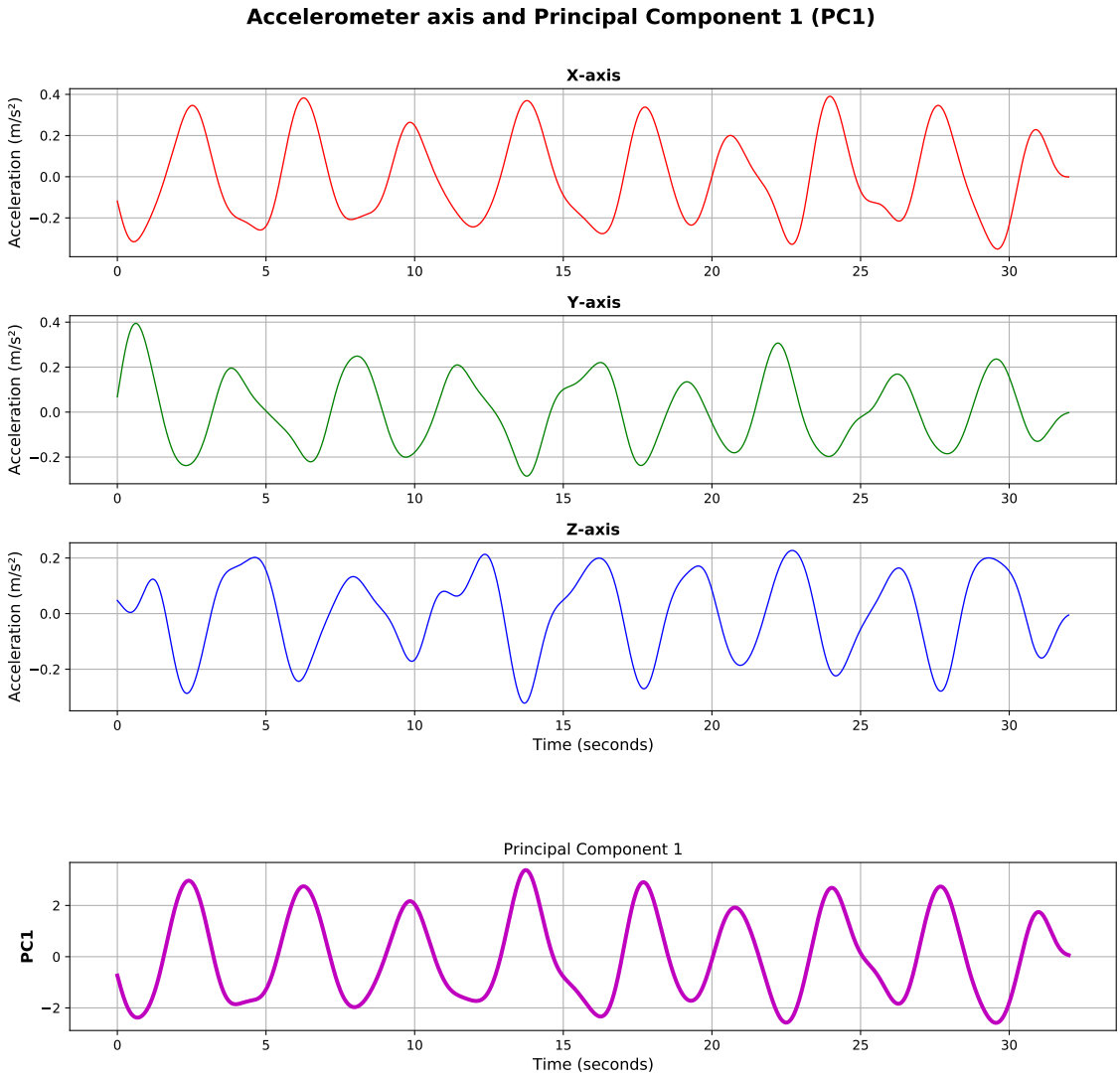

Figure 5.4: Time-series plots of tri-axial ACC data and the PC1. The top three subplots show filtered ACC data along the x, y, and z axes, respectively. The bottom subplot shows the PC1 extracted via PCA, representing the dominant motion pattern across all axes. This visualization highlights the dimensionality reduction from three signals to a single representative signal.

5.4.2 Quality control

Motion-based signals often include a significant amount of noise. Although filtering helps reduce some of this, we also applied additional quality control measures to identify windows with good signal quality for further analysis. We evaluated the same OCSVM method used with PPG as Kazemi *et al.* [9] described, but it produced a very low yield, retaining only 8,4% of the ACC windows. Consequently, we decided to explore other quality control options. Huang *et al.* [8] computed a second-by-second motion activity (MA) level for each window and excluded those in which the MA exceeded a predefined threshold. However, the specific implementation of the MA method was not described in detail, and the threshold is not specified. In addition to MA analysis, they also applied unspecified kurtosis and variance thresholds to further filter out poor-quality windows. Based on this, we decided to utilize three common statistical measures for quality estimation: (1) flat percentage, (2) variance, and (3) kurtosis.

(1) To detect periods of flatness in the signal, we calculate the *flat percentage*, defined as the proportion of rolling windows where the standard deviation falls below a threshold of 0.01. This is done by computing the rolling standard deviation over the signal using a fixed window size of 50 samples. For a given time point t, the rolling standard deviation std_t is defined as:

$$\text{std}_t = \sqrt{\frac{1}{w-1} \sum_{i=t-w+1}^{t} (x_i - \bar{x}_t)^2} \tag{5.5}$$

where w is the window size, x_i are the signal values within the window, and $\bar{x}_t$ is their mean. The denominator $w - 1$ reflects the sample standard deviation, which corrects for the bias in estimating the population variance. The number of windows with standard deviation below the threshold is counted, and the result is expressed as a percentage of all valid windows. This metric helps identify flat segments within

the window which indicate a low-quality signal.

(2) Another way to detect low-quality windows is to calculate the overall *variance* of the window. Variance is a statistical measure of dispersion that quantifies how much the values within a window deviate from their mean. Low variance within a window typically indicates a lack of signal activity or minimal change over time, which may correspond to periods of sensor dropout, inactivity, or flatlining. The variance σ^2 of a window of size w is defined as:

$$\sigma^2 = \frac{1}{w-1}\sum_{i=1}^{w}(x_i - \bar{x})^2 \tag{5.6}$$

where x_i represents individual signal samples within the window, and $\bar{x}$ is the mean of those samples. This method is computationally simple and complements other measures by providing a global view of variability within each segment.

(3) The final quality measure was *kurtosis*, a statistical metric that describes the tailedness or peakedness of a signal's distribution. Kurtosis is defined as:

$$\text{kurtosis} = \frac{1}{w}\sum_{i=1}^{w}\left(\frac{x_i - \bar{x}}{s}\right)^4 \tag{5.7}$$

where x_i are the signal values within a window of size w, $\bar{x}$ is the mean, and s is the standard deviation. A high kurtosis value indicates a distribution with a sharp peak or heavy tails, suggesting the presence of sudden changes, such as movement or noise spikes. In contrast, low kurtosis reflects a flatter, more uniformly distributed signal. Since kurtosis is particularly sensitive to outliers and short-lived artifacts, it can help identify windows that may be distorted by brief, irregular events in the signal. [41]

Determining the thresholds

To assess the quality of each window, we determined thresholds for variance and kurtosis employing a supervised regression approach that models their relationship to the corresponding prediction error. The goal was to identify values of variance and kurtosis beyond which the model error increases significantly, indicating a possible deterioration in signal quality. To avoid overfitting and evaluate model generalization, the dataset was first partitioned into training and test sets using an 80/20 split.

A Random Forest Regressor (with 100 estimators and a fixed random seed for reproducibility) was trained on the training set to learn the mapping from variance and kurtosis to prediction error. Once trained, the model was used to predict errors on the held-out test set. Thresholds were derived from the test set predictions. Specifically, the test data were sorted independently by variance and kurtosis, and the numerical gradient of the predicted error was computed for each. The threshold for each feature was defined as the point corresponding to the maximum gradient—i.e., the feature value at which the predicted error exhibited the most rapid increase. These thresholds reflect the estimated boundary over or under which the model detects a substantial rise in error, and thus can be used as thresholds for quality control. The computed thresholds were 1.308 for variance and 0.038 for kurtosis. For the flat percentage we empirically evaluated different thresholds and found that a threshold of 1% provided a good balance between detecting flat signal segments and preserving high-quality data.

Decision rules

A window was classified as low quality (label 0) if it met all of the following conditions: (1) the flat percentage exceeded the predefined threshold, (2) the variance was below the minimum threshold, and (3) the kurtosis was above the maximum

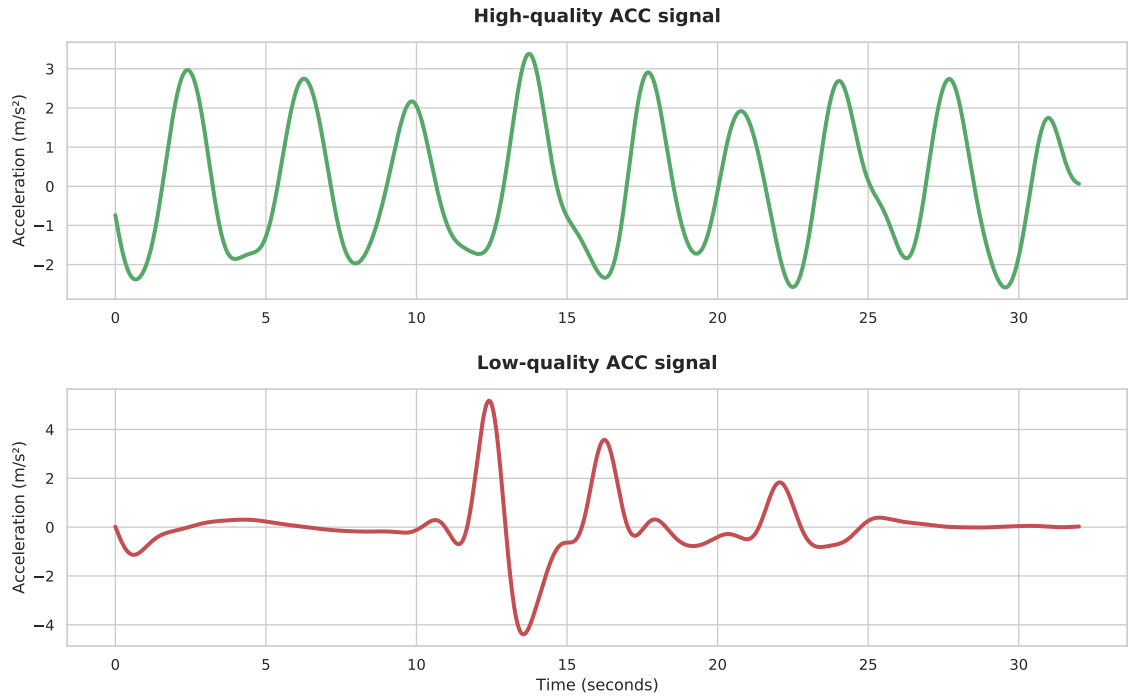

Figure 5.5: Examples of high- and low-quality ACC signals.

threshold. Otherwise, it was labelled as good quality (label 1). The decision rule is defined as:

$$\text{quality}_i = \begin{cases} 0, & \text{if } f_i > \tau_f \text{ and } v_i < \tau_v \text{ and } k_i > \tau_k \\ 1, & \text{if } f_i < \tau_f \text{ and } v_i > \tau_v \text{ and } k_i < \tau_k \end{cases} \tag{5.8}$$

where f_i, v_i, and k_i are the flat percentage, variance, and kurtosis computed for the i-th window, respectively; and τ_f, τ_v, and τ_k are their corresponding thresholds. The difference between high- and low-quality windows can be seen in Figure 5.5. Low-quality windows were excluded from subsequent analysis.

5.4.3 Adaptive RR tracking

To estimate RR from the PC1 of ACC and GYR signals, we employed the method proposed by Huang *et al.* [8]. We decided to use the same approach as the original

pipeline, since the FFT is a widely recognized technique for deriving respiratory signals from motion-based data [5], [23], [34]. Additionally, we investigated whether using the CWT to extract real-time RR could lead to improved estimation accuracy. The algorithm tracks how the RR evolves over time within each 32-second analysis window.

First, an initial RR estimate is obtained using the FFT. The FFT is applied to the PC1 of the signal within the window to obtain its frequency-domain representation. The dominant frequency is identified as the peak with the highest amplitude within the range of 0.13 to 0.66 Hz (corresponding to 8 to 40 BrPM) and is used as the initial RR estimate. Next, the CWT is applied to the same PC1 signal to generate a time-frequency representation. This allows tracking of dominant frequency components as they evolve over time. Each time slice (i.e., each column of the CWT matrix) is analyzed independently. For each slice, the wavelet spectrum is extracted, and a search is performed for a dominant frequency peak within a narrow range ($\pm 10\%$) around the previous RR estimate. If a peak is found, it becomes the updated RR estimate for that time point. If no valid peak is detected, the previous estimate is reused. If no peaks are found across three consecutive time slices, the initial RR estimate from the FFT is used to reset the estimate. After all time slices have been analyzed, the final RR estimate for the window is computed as the average of the RR estimates calculated at each time point.

5.5 Estimate fusion

Final step of the pipeline was to combine the RR estimates from the different modalities. Huang *et al.* [8] and Kazemi *et al.* [9] mentioned that they fused the estimates using the window quality. However, they did not describe the specific implementation, which is why we decided to explore a weighted least squares (WLS) regression model to optimally combine estimates derived from the PPG and ACC signals. The

method was derived from Pimentel *et al.* [35]. The approach leverages the statistical reliability of each modality, derived from the absolute errors relative to the reference data.

We decided to exclude GYR estimates from this last step as the accuracy of the estimates was low. We will discuss this in more detail in Section 6.

5.5.1 Data preprocessing

First we determined the windows where both PPG and ACC signal quality were acceptable and extracted the corresponding RR estimates. In addition to these sets, we extracted the corresponding reference RR values for each window. These RR estimates were combined into tuples and randomly shuffled to prevent any order bias. The dataset was then partitioned into training (80%) and testing (20%) subsets.

5.5.2 Variance-based weight estimation

For each training sample, absolute errors were computed between the modality-specific RR estimates and the reference. The variance of these residuals was then calculated for both PPG and ACC signals. Inverse variances were used to define scalar weights w_{PPG} and w_{ACC} , under the assumption that lower variance corresponds to higher reliability. By taking the inverse of the variance, we give more weight to the modality with the lower variance. The scalar weights are defined as:

$$w_{\text{PPG}} = \frac{1}{\text{Var}(\text{error}_{\text{PPG}})}, \quad w_{\text{ACC}} = \frac{1}{\text{Var}(\text{error}_{\text{ACC}})} \tag{5.9}$$

These scalar weights were used to define sample-specific weights for WLS regression by forming a linear combination of the two RR estimates:

$$\text{weight}_i = w_{\text{PPG}} \cdot \text{RR}_{\text{PPG}i} + w_{\text{ACC}} \cdot \text{RR}_{\text{ACC}i} \tag{5.10}$$

5.5.3 Model fitting and evaluation

The WLS regression model was fit to the training data using PPG and ACC estimates as predictors and the reference RR as the response. The model produced optimal coefficients for combining the two signal modalities (w_1: 0.6174, w_2: 0.2676). The coefficients were evaluated on the held-out test set to assess their performance. The model demonstrated good predictive accuracy, achieving a mean absolute error of 2.367 BrPM. The coefficients were subsequently applied to the full dataset.

5.5.4 Decision rules for combination

Finally, the RR estimates from the two sensor modalities, PPG and ACC, are combined using decision rules, defined as:

$$\hat{RR}_i = \begin{cases} w_1 \cdot \mathrm{RR}_{\mathrm{PPG}i} + w_2 \cdot \mathrm{RR}_{\mathrm{ACC}i} & \text{if } \mathrm{PPG}_i \text{ and } \mathrm{ACC}_i \text{ are good} \\ \mathrm{RR}_{\mathrm{PPG}i} & \text{if only } \mathrm{PPG}_i \text{ is good} \\ \mathrm{RR}_{\mathrm{ACC}i} & \text{if only } \mathrm{ACC}_i \text{ is good} \\ \mathrm{Outage} & \text{if neither is good} \end{cases} \tag{5.11}$$

where, $\hat{RR}_i$ denotes the final RR estimate for the i^{th} sample. The variables $\mathrm{RR}_{\mathrm{PPG}i}$ and $\mathrm{RR}_{\mathrm{ACC}i}$ represent the individual estimates from the PPG and ACC modalities, respectively. The scalar weights w_1 and w_2 are the coefficients derived from the WLS model.

The selection logic prioritizes the use of both signals when both are deemed reliable (i.e., pass the quality thresholds). If only one modality passes quality control, its estimate is used directly. Samples for which neither signal meets quality criteria are determined to be outages in the analysis.

6 Testing and results

To evaluate the performance of the proposed pipeline, we used the wrist-worn signals (PPG, ACC, GYR) alongside the reference respiratory signal collected from 26 of the 45 individuals. To maximize data reliability, we concentrated on recordings collected during the night when noise levels are typically lower. 19 participants were excluded from the analysis due to technical issues that resulted in unsynchronized data.

We wanted to answer three questions with the proposed pipeline: (1) can it produce accurate RR estimates using PPG and motion-based data (ACC or GYR), (2) can the pipeline be constructed using only waveform analysis methods, and (3) does adding the motion-based estimates enhance the accuracy of RR estimates derived from PPG signals. We used the pipeline proposed by Huang *et al.* [8] and tested by Kazemi *et al.* [9] as a baseline and tried to improve its accuracy. The final pipeline architecture is illustrated in Fig. 6.2.

6.1 Signal quality assessment

The proportion of time windows classified as high quality varied across sensor modalities. The analysis of signal quality revealed that 45% of ACC windows, 56% of GYR windows, and 77% of PPG windows met the quality criteria defined in the Sections 5.4.2 and 5.3.1, highlighting PPG as the most reliable modality.

6.2 Performance measures

First, the estimation error for each method was quantified by comparing the predicted RR in each 32-second analysis window to the corresponding reference value. The error for each window was computed as:

$$\text{Error}_i = \hat{RR}_i - RR_i^{\text{ref}} \tag{6.1}$$

where $\hat{RR}_i$ is the method-predicted RR and RR_i^{ref} is the corresponding reference value for window i. Positive errors indicate overestimation, while negative errors indicate underestimation.

To evaluate the performance of each method, three standard error metrics were calculated: mean absolute error (MAE), root mean square error (RMSE), and mean bias [9]. Let $\hat{RR}_i$ denote the estimated RR and RR_i^{ref} the corresponding reference value for the i^{th} analysis window, with a total of N valid samples.

MAE measures the average magnitude of the estimation error, regardless of direction. It provides a good measure of overall estimation accuracy:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{RR}_i - RR_i^{\text{ref}}\right| \tag{6.2}$$

RMSE emphasizes larger errors by squaring the differences before averaging, and is sensitive to outliers. It is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{RR}_i - RR_i^{\text{ref}}\right)^2} \tag{6.3}$$

Mean Bias measures the average signed error, indicating whether a method tends to systematically overestimate or underestimate the true value:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{RR}_i - RR_i^{\text{ref}}\right) \tag{6.4}$$

To quantify the uncertainty of the MAE and RMSE estimates, 95% confidence intervals were calculated using the standard error of the mean, based on the sample standard deviation:

$$\mathrm{CI}_{95\%} = 1.96 \times \frac{s}{\sqrt{N}} \tag{6.5}$$

where s is the sample standard deviation of the absolute errors (for MAE) or the squared errors (for RMSE).

6.3 Performance results of the methods

Testing the adaptive RR tracking method with ACC data revealed that incorporating CWT did not improve the RR estimates; in fact, it yielded results identical to those obtained using the dominant frequency derived from FFT alone. Consequently, we opted to use only the FFT-based approach for the subsequent analysis of ACC and GYR signals. This improves the efficiency of the pipeline, as we do not need to compute the CWT of the ACC or GYR signals.

Table 6.1 presents the performance metrics for the evaluated RR estimation methods across different modalities. The modulation-based approach using PPG achieved the best overall performance, achieving the lowest MAE (2.425 ± 0.051 BrPM) and RMSE (3.007 ± 0.360 BrPM) with yield of 77%, along with a low mean bias of 0.040 BrPM. This approach outperformed the CWT-based estimation and was therefore selected to be used in the fusion method. Estimates based on the ACC and GYR signals resulted in significantly higher MAE and RMSE values, with GYR performing the worst (MAE: 4.964 ± 0.002 BrPM and RMSE: 6.551 ± 0.047 BrPM). Due to the low performance of the GYR analysis, we decided to exclude it from the fusion method. The fusion method, which combines estimates from PPG and ACC, performed similarly compared to the modulation-based approach using

PPG, with an MAE of 2.492 ± 0.052 BrPM, RMSE of 3.132 ± 0.409 BrPM, and a mean bias of 0.067 BrPM. Notably, the fusion method achieved an increase in yield from 77% to 82%. The method also outperformed the similar pipeline Kazemi *et al.* [9] tested using the same dataset utilized in this thesis, with an MAE of 2.90 ± 0.70 BrPM, RMSE of 3.96 ± 0.60 BrPM and a mean bias of -2.2 BrPM. However, the CNN-based method Kazemi *et al.* [9] proposed, performed significantly better (MAE: 1.85 ± 0.40 BrPM, RMSE: 2.34 ± 0.30 BrPM). Notably, they did not report the yield of the corrupted segment elimination process.

Table 6.1: Performance metrics for RR estimation methods across different modalities. MAE and RMSE are reported as estimate ± 95% confidence interval.

Signal	Method	MAE (BrPM)	RMSE (BrPM)	Mean Bias (BrPM)
ACC	FFT	3.992 ± 0.121	5.136 ± 1.666	3.399
GYR	FFT	4.964 ± 0.002	6.551 ± 0.047	4.414
PPG	CWT	2.559 ± 0.056	3.219 ± 0.435	-0.738
PPG	Modul.	2.425 ± 0.051	3.007 ± 0.360	0.040
PPG + ACC	Modul. + FFT	2.492 ± 0.052	3.132 ± 0.481	0.067

To further assess the performance of each method, we examined the distribution of estimation errors using box plots (Figure 6.1), which provide a detailed view of the median error, variability, and the presence of outliers for each estimation approach. The modulation-based PPG RR estimation and fusion method achieved near-zero median absolute errors (-0.07 and -0.06 BrPM, respectively), suggesting a strong agreement with the reference RR. The methods relying on ACC and GYR exhibited significantly positive median errors (3.04 and 3.88 BrPM, respectively), indicating a consistent overestimation of RR relative to the reference. The GYR analysis resulted in a substantial number and range of outliers, suggesting that the method is less robust and more prone to failure under atypical conditions. It is also

worth noting that the fusion approach results in more outliers than the modulation-based PPG RR estimation, implying that fusion of the estimates does not improve the estimation.

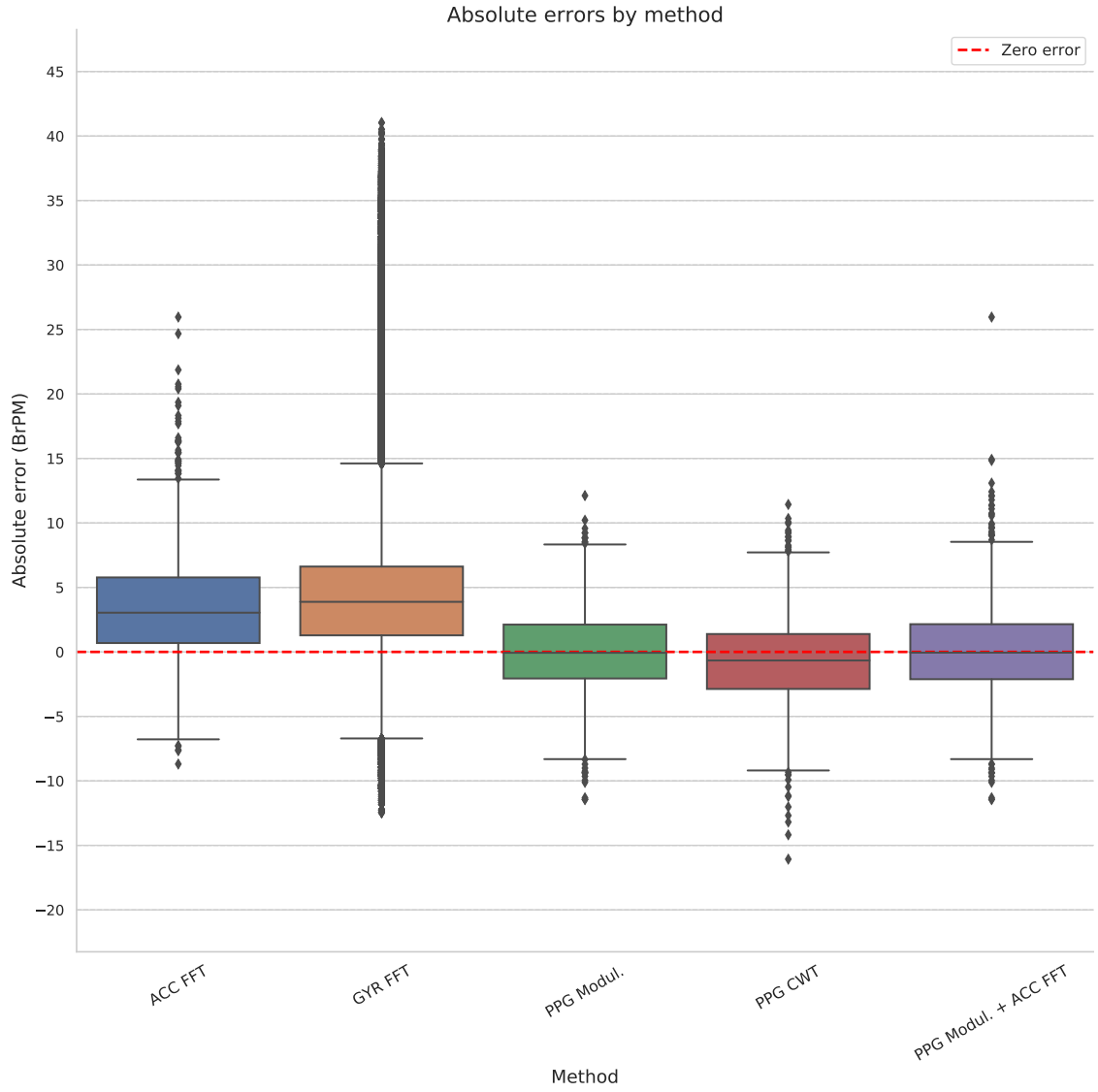

Figure 6.1: Box plot of the absolute RR estimation errors for each method: ACC (FFT), GYR (FFT), PPG (Modulation-based estimation), PPG (CWT), and the Fusion approach (PPG modul. + ACC FFT). Each box illustrates the interquartile range (IQR), capturing the middle 50% of the errors, while the horizontal line within each box represents the median absolute error. Whiskers extend to 1.5 times the IQR, with points beyond are considered outliers. A dashed red line at zero BrPM indicates perfect estimation accuracy.

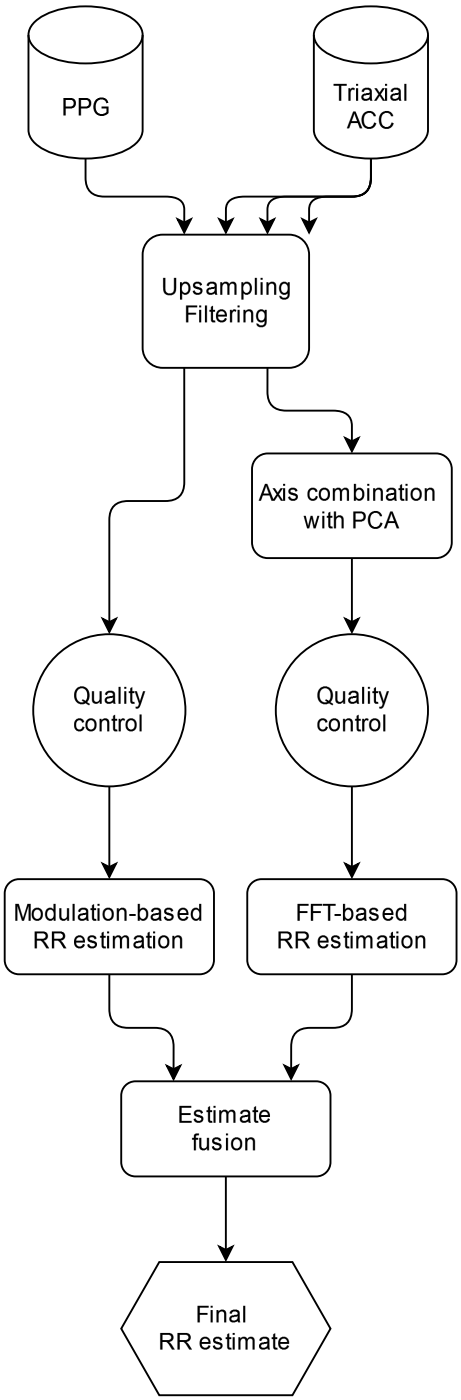

Figure 6.2: Visualisation of the final pipeline architecture.

7 Discussion

This thesis set out to examine three research questions: RQ1: Can an algorithm that integrates PPG and motion-based data produce accurate RR estimates? RQ2: Can this be done by only using waveform analysis methods? RQ3: Does incorporating motion-based RR estimates enhance the accuracy of RR estimation derived from PPG signals? The results provide varying levels of support for each: RQ1 was confirmed, RQ2 was partially supported, and RQ3 was not supported.

RQ1 was confirmed as we proved that an fusion method that uses PPG and motion-based signals (ACC in this case) produces fairly accurate RR estimates with a minimal mean bias of nearly zero. RQ2 was partially confirmed, as we exclusively employed waveform analysis techniques to extract the respiratory-related components. These methods included frequency-domain analysis using FFT to identify the dominant frequency in the ACC signals, time-domain analysis to capture respiration-induced modulation in the PPG signals, and for comparison time-frequency analysis of PPG signals using the CWT. However, we utilized various machine learning techniques for signal quality control. We determined the ACC signal quality thresholds with a random forest regressor which was chosen as it was difficult to find optimal thresholds experimentally due to high level of noise in the motion-based signals. For PPG quality control, we chose the OCSVM method developed in [28] due to its demonstrated accuracy and efficiency. In conclusion, RR extraction was achieved through waveform analysis techniques, complemented by machine learning for signal

quality control. RQ3 was not supported, as the fusion method did not outperform the modulation-based estimation using only PPG. However, its performance was only marginally lower and the fusion increased the amount of valid windows from 77% to 82%. Given the added complexity of the fusion method's processing pipeline, the single-modality modulation-based RR estimation from PPG proves to be the more efficient and accurate choice.

Additionally, we succeeded in improving the accuracy of the pipeline compared to the original pipeline proposed by Huang *et al.* [8] and tested by Kazemi *et al.* [9]. Although multiple methods were evaluated throughout the pipeline development, we ultimately selected those that demonstrated the highest accuracy. As a result, the final fusion pipeline closely resembles the original structure. Notably, to streamline the ACC processing, the CWT step was removed as it provided no measurable benefit. More importantly, the GYR signals were excluded from the pipeline altogether due to their consistently low estimation accuracy. This modification is likely a key factor behind the improvement in pipeline accuracy.

One of the most challenging aspects of the analysis was assessing and improving the quality of the ACC and GYR signals, which were clearly heavily affected by noise. Respiration-related movements in these signals are subtle and can be easily overpowered by the other types of motion captured by the sensors. Therefore, improving noise reduction techniques for motion-based signals is crucial for increasing the proportion of high-quality data windows. The improvements could enhance signal reliability but also contribute to the overall accuracy of the fusion method, as RR estimates derived from ACC exhibited considerably higher error rates compared to those obtained from PPG.

References

[1] M. Filippelli, R. Pellegrino, I. Iandelli, *et al.*, "Respiratory dynamics during laughter", *Journal of Applied Physiology*, vol. 90, no. 4, pp. 1441–1446, 2001.

[2] S. Fuchs and A. Rochet-Capellan, "The respiratory foundations of spoken language", *Annual Review of Linguistics*, vol. 7, pp. 13–30, 2021.

[3] M. A. Cretikos, R. Bellomo, K. Hillman, J. Chen, S. Finfer, and A. Flabouris, "Respiratory rate: The neglected vital sign", *Medical Journal of Australia*, vol. 188, no. 11, pp. 657–659, 2008.

[4] T. Hao, C. Bi, G. Xing, R. Chan, and L. Tu, "Mindfulwatch: A smartwatch-based system for real-time respiration monitoring during meditation", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–19, 2017.

[5] J. Hernandez, D. McDuff, and R. W. Picard, "Biowatch: Estimation of heart and breathing rates from wrist motions", in *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, IEEE, 2015, pp. 169–176.

[6] D. Liaqat, M. Abdalla, P. Abed-Esfahani, *et al.*, "Wearbreathing: Real world respiratory rate monitoring using smartwatches", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 2, pp. 1–22, 2019.

[7] R. Dai, C. Lu, M. Avidan, and T. Kannampallil, "Respwatch: Robust measurement of respiratory rate on smartwatches with photoplethysmography", in *Proceedings of the international conference on internet-of-things design and implementation*, 2021, pp. 208–220.

[8] N. Huang, M. Zhou, D. Bian, *et al.*, "Novel continuous respiratory rate monitoring using an armband wearable sensor", in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2021, pp. 7470–7475.

[9] K. Kazemi, I. Azimi, P. Liljeberg, and A. M. Rahmani, "Robust cnn-based respiration rate estimation for smartwatch ppg and imu", in *Proceedings of the 2023 10th International Conference on Bioinformatics Research and Applications*, 2023, pp. 94–100.

[10] P. B. Lovett, J. M. Buchwald, K. Stürmann, and P. Bijur, "The vexatious vital: Neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage", *Annals of emergency medicine*, vol. 45, no. 1, pp. 68–76, 2005.

[11] E. Kaniusas, "Fundamentals of biosignals", in *Biomedical signals and sensors I*, Springer, 2012, pp. 1–26.

[12] F. Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan, "Respiration rate monitoring methods: A review", *Pediatric pulmonology*, vol. 46, no. 6, pp. 523–529, 2011.

[13] C. Subbe, R. Davies, E. Williams, P. Rutherford, and L. Gemmell, "Effect of introducing the modified early warning score on clinical outcomes, cardiopulmonary arrests and intensive care utilisation in acute medical admissions", *Anaesthesia*, vol. 58, no. 8, pp. 797–802, 2003.

[14] M.-C. Huang, W. Xu, J. Liu, *et al.*, "Inconspicuous on-bed respiratory rate monitoring", in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments*, 2013, pp. 1–8.

[15] R. Gilbert, J. Auchincloss Jr, J. Brodsky, and W. a. Boden, "Changes in tidal volume, frequency, and ventilation induced by their measurement.", *Journal of Applied Physiology*, vol. 33, no. 2, pp. 252–254, 1972.

[16] P. A. Dargaville, I. Frerichs, D. Tingay, *et al.*, "Monitoring lung volumes during mechanical ventilation", in *Pediatric and Neonatal Mechanical Ventilation: From Basics to Clinical Practice*, Springer, 2014, pp. 441–471.

[17] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care", *International journal of biosensors & bioelectronics*, vol. 4, no. 4, p. 195, 2018.

[18] D. J. Meredith, D. Clifton, P. Charlton, J. Brooks, C. Pugh, and L. Tarassenko, "Photoplethysmographic derivation of respiratory rate: A review of relevant physiology", *Journal of medical engineering & technology*, vol. 36, no. 1, pp. 1–7, 2012.

[19] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram", *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1946–1953, 2013.

[20] J. Li, J. Jin, X. Chen, W. Sun, and P. Guo, "Comparison of respiratory-induced variations in photoplethysmographic signals", *Physiological measurement*, vol. 31, no. 3, p. 415, 2010.

[21] J. Kim, J. Kim, and H. Ko, "Low-power photoplethysmogram acquisition integrated circuit with robust light interference compensation", *Sensors*, vol. 16, no. 1, p. 46, 2015.

[22] P. Dehkordi, A. Garde, B. Molavi, J. M. Ansermino, and G. A. Dumont, "Extracting instantaneous respiratory rate from multiple photoplethysmogram respiratory-induced variations", *Frontiers in physiology*, vol. 9, p. 948, 2018.

[23] M. Haescher, D. J. Matthies, J. Trimpop, and B. Urban, "A study on measuring heart-and respiration-rate via wrist-worn accelerometer-based seismocardiography (scg) in comparison to commonly applied technologies", in *Proceedings of the 2nd international Workshop on Sensor-based Activity Recognition and Interaction*, 2015, pp. 1–6.

[24] V. M. Passaro, A. Cuccovillo, L. Vaiani, M. De Carlo, and C. E. Campanella, "Gyroscope technology and applications: A review in the industrial perspective", *Sensors*, vol. 17, no. 10, p. 2284, 2017.

[25] A. Schäfer and K. W. Kratky, "Estimation of breathing rate from respiratory sinus arrhythmia: Comparison of various methods", *Annals of Biomedical Engineering*, vol. 36, pp. 476–485, 2008.

[26] S. Butterworth *et al.*, "On the theory of filter amplifiers", *Wireless Engineer*, vol. 7, no. 6, pp. 536–541, 1930.

[27] C. Orphanidou, S. Fleming, S. A. Shah, and L. Tarassenko, "Data fusion for estimating respiratory rate from a single-lead ecg", *Biomedical Signal Processing and Control*, vol. 8, no. 1, pp. 98–105, 2013.

[28] M. Feli, I. Azimi, A. Anzanpour, A. M. Rahmani, and P. Liljeberg, "An energy-efficient semi-supervised approach for on-device photoplethysmogram signal quality assessment", *Smart Health*, vol. 28, p. 100 390, 2023.

[29] M. A. Pimentel, A. E. Johnson, P. H. Charlton, *et al.*, "Toward a robust estimation of respiratory rate from pulse oximeters", *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1914–1923, 2016.

[30] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series", *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

[31] S. D. Tang, Y. S. Goh, M. D. Wong, and Y. E. Lew, "Ppg signal reconstruction using a combination of discrete wavelet transform and empirical mode decomposition", in *2016 6th International Conference on Intelligent and Advanced Systems (ICIAS)*, IEEE, 2016, pp. 1–4.

[32] A. M. Sururi, E. S. Pramukantoro, K. Amron, V. Wardhani, and P. A. Kamila, "Comparison of several wavelet transform techniques for inferring ecg signal classification", in *Proceedings of the 8th International Conference on Sustainable Information Engineering and Technology*, 2023, pp. 167–172.

[33] J. Leube, J. Zschocke, M. Kluge, *et al.*, "Reconstruction of the respiratory signal through ecg and wrist accelerometer data", *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.

[34] T. Reinvuo, M. Hannula, H. Sorvoja, E. Alasaarela, and R. Myllyla, "Measurement of respiratory rate with high-resolution accelerometer and emfit pressure sensor", in *Proceedings of the 2006 IEEE Sensors Applications Symposium, 2006.*, IEEE, 2006, pp. 192–195.

[35] M. A. Pimentel, P. H. Charlton, and D. A. Clifton, "Probabilistic estimation of respiratory rate from wearable sensors", in *Wearable electronics sensors*, Springer, 2015, pp. 241–262.

[36] M. A. Mehrabadi, I. Azimi, F. Sarhaddi, *et al.*, "Sleep tracking of a commercially available smart ring and smartwatch against medical-grade actigraphy in everyday settings: Instrument validation study", *JMIR mHealth and uHealth*, vol. 8, no. 11, e20465, 2020.

[37] W. Karlen, J. M. Ansermino, and G. Dumont, "Adaptive pulse segmentation and artifact detection in photoplethysmography for mobile applications", in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2012, pp. 3131–3134.

[38] F. Schipper, R. J. van Sloun, A. Grassi, R. Derkx, S. Overeem, and P. Fonseca, "Estimation of respiratory rate and effort from a chest-worn accelerometer using constrained and recursive principal component analysis", *Physiological Measurement*, vol. 42, no. 4, p. 045 004, 2021.

[39] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling", *IEEE Transactions on medical imaging*, vol. 2, no. 1, pp. 31–39, 1983.

[40] M. Elgendi, I. Norton, M. Brearley, D. Abbott, and D. Schuurmans, "Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions", *PloS one*, vol. 8, no. 10, e76585, 2013.

[41] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in photoplethysmogram", in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2011, pp. 4972–4975.