Applied Computer Science, vol. 20, no. 1, pp. 139–156
doi: 10.35784/acs-2024-09

Submitted: 2023-06-11 | Revised: 2024-02-12 | Accepted: 2024-02-21

Keywords: signal processing, digital filters, sensors, technology

Alexandru Marius OBRETIN [0009-0001-0952-9109]*,
Andreea-Alina CORNEA [0009-0004-4373-3949]*

FILTERING STRATEGIES FOR SMARTPHONE EMITTED DIGITAL SIGNALS

Abstract

In today's digitalized and technology-driven society, where the number of IoT devices and the volume of collected data is exponentially increasing, the use of sensor data has become a necessity in certain fields of activity. This paper presents a concise history of sensor evolution and specialization, with a focus on the sensors used for localization, particularly those present in microelectromechanical systems (MEMS) that make up inertial measurement units. The study starts with a general overview and progresses towards a more specific analysis of data sets collected from an accelerometer. In the materials and methods section, it emphasizes the imperfect nature of sensor measurements and the need to filter digital signals. Three different digital signal filtering techniques belonging to distinct filter categories are comparatively analyzed, with each technique having its own particularities, advantages and disadvantages. The analysis considers the effectiveness in reducing measurement errors, the impact of the filtering process on the original signal, the ability to highlight the underlying phenomenon, as well as the performance of the analyzed algorithms. The ultimate purpose of this article is to determine which filtration method is best suited for the signal at hand in the context of an indoor localization application.

1. INTRODUCTION

The concept of sensors is tightly coupled to the evolution of technology and the reinterpretation of daily activities. They are widely used in numerous contexts, both implicitly and explicitly, ranging from the medical field, where they are essential for monitoring patients' vital signs, to domestic settings, where they are integrated into mobile phones and other electronic devices to identify location, measure environmental parameters and even gauge a person's current state, physical activity progress and other relevant indicators of effort or sleep quality, having these functionalities offered by the emergence of smart components in recent years. Despite being predominantly associated with the digital revolution, the history of sensors or their precursors can be traced back several centuries to the works of Leonardo da Vinci, who developed temperature and pressure measurement techniques, and Jacques Curie's documentation of materials that can generate voltage. The

* Bucharest University of Economic Studies, Faculty of Cybernetics, Statistics and Economic Informatics, Department of Economic Informatics, Romania, alexandru.obretin@csie.ase.ro, andreea.cornea@csie.ase.ro

use of sensors gained further traction during the World Wars when they were employed in the detection of enemy elements through radar or sonar. This historical background laid the foundation for the development of components used across various industries to gather crucial data for proper operation. The literature identifies practical use cases for sensors in fields such as food, medical, automotive, air, and military, among others, highlighting their indispensable role in facilitating the functioning of the economy.

Sensors constitute a fundamental component of data extraction in the food industry, enabling the implementation of Big Data concepts by facilitating efficient management of large volumes of data (Dinakar & Vagdevi, 2017). The information obtained through sensors is integrated into all stages of production, processing, distribution, transport and marketing. The accuracy of this information ensures industry workers can rely on it to implement processes that maintain business continuity and comply with regulations imposed by the legislation in force. Sensors provide an easy way to obtain environmental details during the stages mentioned above. For example, they can provide information on temperature throughout the entire product chain management process to guarantee food safety. Sensors can also monitor agricultural growth levels, directing workers towards production that is ready for harvest and can be sent forward in accordance with the predefined flow of operations.

In the medical field, sensors find diverse applications, ranging from monitoring vital functions in various settings such as the hospitals or even household devices (Mohamed et al., 2023). This wide applicability enables the detection of anomalies and the diagnosis of diseases (Ahlawat et al., 2024). Moreover, sensors are integrated into tools used for imaging purposes, where image sensors provide visual differences between healthy and affected surfaces, aiding in overall analysis (Yadav et al., 2024). Notably, sensors also play a critical role in treatments, with applications in implantable devices and providing support for monitoring conditions (Xie et al., 2017).

In the automotive industry, technological advancements have revolutionized the way cars are viewed and used, offering a more enjoyable and safer driving experience. Sensors are a modern approach to monitoring vehicle conditions, improving performance and increased safety (Bloecher et al., 2009; Karpinski et al., 2023). They support driver-assistance systems, enabling braking and parking assistance while constantly detecting obstacles, thus avoiding possible accidents and providing a personalized experience. The integration of sensors into vehicle operation is not a recent implementation but dates back to the past, having such components used to monitor liquid temperature or tire pressure. With technological advancements, the automotive industry has implemented new functionalities to attract customers, including a friendlier relationship with the environment, through sensors that measure discharged polluting substances.

In the fields of aerospace and military technology, the use of state-of-the-art and accurate devices is imperative to ensure the highest level of safety (Azam et al., 2023; Anwar et al., 2023). The Inertial Measurement Unit (IMU) sensor, which consists of an accelerometer, gyroscope and magnetometer, is particularly crucial for both domains, as it enables the orientation of an object in space and the determination of its direction of movement (Gujarathi & Bhole, 2019).

In the aviation industry, the IMU plays a crucial role in ensuring safe flight management and is an extension of the pilot. The IMU monitors the speed, orientation, and position of the aircraft in space, enabling it to continue transport under adverse environmental conditions and maintain a safe distance from other obstacles in the airspace. Additionally, just as sensors are used in the automotive field for driving assistance systems, they can also be integrated into aircraft to facilitate flights with automatic pilot or assistance during take-offs, landings, emergencies or unforeseen situations, relying on accurate navigation data to calculate the plane's position and determine the risks of each possible trajectory (Zhang et al., 2005).

In the realm of military operations, similar to the airfield, the IMU serves multiple purposes, intersecting with previous examples in military aviation. Additionally, it can be used in land or sea operations to track vehicle movement, monitor devices, ensure movement in adverse weather or environmental conditions and detect unknown, potentially hostile objects that may be situated near devices monitored by IMU sensors. To comprehensively understand the IMU's mode of operation, it is necessary to synthesize the functionalities of each integrated sensor at the assembly level.

The accelerometer is used to calculate acceleration and is based on inertial forces to determine an object's speed of movement, by measuring movement on the tridimensional axis system and converting it into electrical signals, which then transmits to the component to which it is integrated (Promrit et al., 2018).

Similarly, the gyroscope is based on inertial forces, aiming to determine the orientation and rotation of an object, and is represented by a rotor that goes around an axis. The gyroscopic effect is also used at the gyroscope level to allow for the identification of orientation and the preservation of direction in the situation where the sensor is in motion, in space, thus determining the angle of displacement (Olivares et al., 2009).

The magnetometer is based on the magnetic properties of materials and their interaction in space, with the sensor's functionality provided by the measurement of magnetic fields. Based on this principle, the position of the device containing this sensor can be determined in relation to magnetic North, due to the induction of electric voltage in the direction of the magnetic field in which the device is oriented. Practically, in the context of using a mobile phone, the magnetometer is based on the Hall Effect, which generates a magnetic field when a magnet is placed near the device, thus determining the orientation (Poulose et al., 2019).

In the domain of mobile devices, a set of components known as microelectromechanical systems (MEMS) are combined to provide the necessary sensors for the phone's key functions. The MEMS include the accelerometer, magnetometer and gyroscope, as well as pressure, light, and proximity sensors. They allow the development of complex functionalities and enable users to access intelligent applications that extract data from the sensors (Sung et al., 2014).

Given the increasing reliance on mobile phones for daily activities, such as making phone calls, accessing social media, determining distance, and measuring environmental factors like temperature, it is imperative that the information provided by the phone is accurate. Consequently, the precision offered by these sensors has been evaluated and found to have measurement errors, requiring correction methods to improve accuracy of the application.

Even under static conditions, where the device is placed on a flat surface and exposed to no external forces except for gravity, the sensors in smart mobile devices record measurement errors. These errors can be attributed to manufacturing defects, physical limitations imposed by the device specifications, calibration errors, physical damage to the device, software component errors responsible for processing signals collected from the sensors, improper use of the device during measurements, environmental factors such as humidity, pressure, temperature or magnetic field interference and the degree of wear and charge of the battery or external electrical noise, which can interfere with the measured phenomenon (Truzman et al., 2021; Bentler & Chiou, 2006; Tan & Jiang, 2018).

It is imperative to consider all the factors that may influence the collection and measurement of digital signals. Despite taking meticulous precautions to minimize the impact of external factors, the measurements obtained from sensors, such as those found in smartphones, are susceptible to noise and errors. Moreover, these measurements tend to accumulate offset over time, leading to a drifting state. Therefore, it is essential to implement a calibration process to ensure the accuracy of the collected data (Vaseghi, 2008).

For illustrative purposes, this article will focus on a 3-axis accelerometer, which is commonly found in mobile phones. Fig. 1 graphically depicts the related triaxle system, which will be referred to throughout this article, when discussing acceleration values on specific axes.

Fig. 1. 3D coordinates system relative to a mobile device

Offset refers to the non-zero measurement value that a sensor indicates even in the absence of acceleration. In the case of a smartphone placed horizontally on a flat surface, ignoring the effects of gravity, the values related to the OX and OY axes should ideally be zero. However, in such circumstances, the measured values may deviate from zero, as shown in Tab. 1.

Tab. 1. Accelerometer values collected in a static context

Timestamp	Roll (X)	Pitch (Y)	Yaw (Z)
1677182050544	-0.6987	-1.4534	-9.8387
1677182050561	-0.0007	-0.7920	-10.7154
1677182050578	-0.1232	-0.7684	-10.2666
1677182050595	-0.5406	-0.9493	-9.6350

The phenomenon of offset, which is also known as bias, is attributable to a variety of sources, including manufacturing defects, temperature fluctuations and external environmental factors. Offset can cause the sensor to produce inaccurate or imbalanced readings which, if not addressed correctly, can significantly compromise the intended outcomes. To mitigate the impact of offset, it is typically advisable to calibrate the accelerometer and eliminate its values by applying correction factors.

Drift is a gradual and unintended shift in the readings of the accelerometer over time. It results in a slight, constant tendency to deviate, even when the user is moving in a straight line, which can lead to readings that indicate a left or right deviation from the actual path. Drift is determined by factors such as mechanical stress, electrical noise and temperature changes, and it can considerably affect the accuracy of the accelerometer readings, even causing it to depart from the reference plane on small surfaces. Mechanical stabilization techniques and signal processing algorithms are employed to minimize the drift tendency of the accelerometer. The subsequent chapter is focused on the use of correction filters to this end.

2. MATERIALS AND METHODS

Digital filters are systems designed to process digital signals through the manipulation of the amplitude and phase of the signal at various frequencies, by using sets of mathematical operations. The fundamental principle of digital filtering involves the convolution of the input signal with a kernel filter - a set of coefficients that define the frequency response of the filter.

There are various ways to classify digital filters, but they will be categorized based on the type of impulse response into finite impulse response (FIR) filters and infinite impulse response (IIR) filters. FIR filters convolve the input signal with a finite number of coefficients, producing a weighted sum of past and present samples of the input signal. The name finite is used to denote that the impulse response of these filters reaches zero within a finite time interval. These properties of FIR filters make them inherently stable, in contrast to IIR filters, where the presence of feedback can impact stability. FIR filters are typically employed when linear phase characteristics are desired, such as in audio equalization or digital signal processing.

On the other hand, IIR filters use feedback to generate an impulse response that extends infinitely in time. The output of an IIR filter is created based on both the current and past samples of the input signal, as well as the previous results of the filtering process. This feedback component represents the essential part of filtering in IIR filters. IIR filters are generally used when a sharper frequency response, with a narrow transition band and sharp cutoff, is required.

2.1. Simple moving average filter

One of the most widely known and effective filtering techniques is the moving average filter, which is a type of signal processing filter that utilizes the average of a set of neighboring values to smooth the overall representation of a wave and remove high and low frequency noise. The algorithm for a simple moving average filter involves operating on a data collection of size M and a predetermined size N, where the mathematical relationship M greater or equal than N must hold. The first filtered value is obtained by calculating the arithmetic mean of the first N terms in the collection. Starting with the second filtered value, the reference subset of size N is shifted to the upper bound of the data set, excluding the first element of the series and including the next available instance. Based on the new subset, the average of the observations in the window is determined, resulting in a smoothed version of the original signal, with the degree of smoothing being determined by the size of the window.

This type of filter is commonly used in conjunction with time series datasets to eliminate short-term fluctuations and emphasize long-term trends or cycles. The distinction between short-term and long-term may differ depending on the scenario and the parameters of the moving average will be adjusted accordingly. The literature suggests that, when utilizing a moving average filter, the window size, impact on the limits, frequency response, computational complexity, and delay induced in the processed signal and inability to maintain abrupt signal transitions should all be considered (Johnston et al., 1999; Hansun & Kristanda, 2017; Ellis & Parbery, 2005).

Despite its limitations, such as its inability to preserve sudden transitions in the signal, this filter is frequently used due to its simplicity, effectiveness in reasonable data volumes and ease of implementation and comprehension (Hidayat et al., 2015; Thinh et al., 2018; Purnama et al., 2022; Ahn & Ko, 2009). The time complexity of this algorithm is linear and it is a robust and effective noise removal tool, with a reduced sensitivity to data anomalies and sudden transitions. By changing the window size or using variations such as weighted, exponential or adaptive averages, this filter becomes a highly flexible tool for filtering digital signals.

To map displacement using an inertial measurement unit, three sets of accelerometer data were collected, each containing acceleration values for all three axes. For ease of comprehension, subsequent plots were generated using a subset of Y-axis acceleration values. To identify the most effective filtering technique, multiple instances of the simple moving average filter were empirically tested using different window sizes.

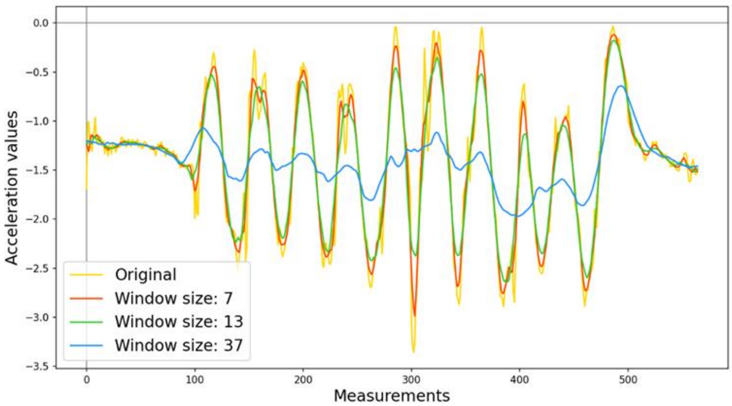

Fig. 2. Simple moving average filter run on different window sizes

Fig. 2 illustrates the recorded accelerometer values and the results of the filtering process for three cases, selected based on the representativeness criteria. The figure demonstrates that a filter with a window size of 7 is insufficient in eliminating all accelerometer measurement errors, as it graphically highlights sinusoids with multiple local extrema points. In contrast, a filter with a window size of 37 elements fails to accurately capture the movement trend. For the analyzed subset, comprising ten steps surrounded by rest periods, the transition between steps 5 to 6 is not clearly highlighted and the local extrema points are difficult to identify, potentially leading to the exclusion of some genuine steps, such as steps 8 and 9.

However, using a filter with a window size of 13 units produces excellent results in terms of accurately and precisely identifying individual steps and smoothing secondary extrema points. Comparable results can be obtained for window sizes close to 13, but the observed differences do not justify the selection of another value over 13.

The choice of the appropriate window size was not only based on graphical results obtained through successive runs, but also on a performance criterion. To this end, a comparison was conducted by running 10,000 simulations for the scenario described earlier, using different window sizes of a simple moving average filter, namely: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, and 37.

The results are presented in Fig. 3, with the horizontal axis representing the window sizes and the vertical axis representing the average processing time in seconds. The range of measurements obtained was between 0.0646 and 0.07827 seconds, with an outlier value of 0.12163 seconds, which was eliminated due to the temporary load of the microprocessor and RAM memory of the used computer system, rather than any particularity of the value 11.

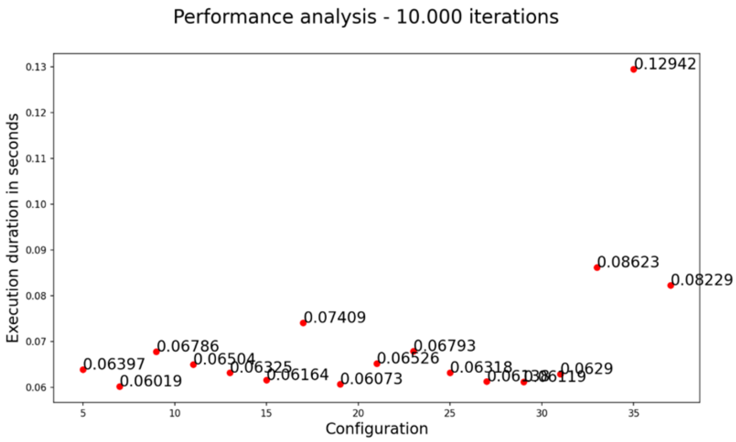

Fig. 3. Performance analysis for the simple moving average filters presented above

Based on a dataset of 565 points, it was observed that the obtained values tended to fall within the same range. Therefore, the decision to choose a certain value depended primarily on the fidelity of the graphical identification of the steps. However, when the dataset was increased experimentally to 307,360 points, the average execution time increased to a range between 22.11978 and 26.83928 seconds. This comparison highlights the importance of implementing an approach based on processing windows, where small volumes of data are analyzed point by point to avoid overloading computing resources and to determine filtered values almost in real time.

2.2. Savitzky-Golay filter

The Savitzky-Golay filter is a sophisticated technique specifically designed for processing digital signals. It is a type of digital smoothing filter that is commonly used to improve the accuracy of a dataset without altering or degrading its underlying structure. This process involves fitting a polynomial curve to a sample of neighboring observations and determining an estimated value for the smoothed sample.

The filter is highly effective at removing noise from the spectrum of processed signals, while preserving important characteristics of the signal such as the presence and location of local extrema points. To achieve this, the filter fits a polynomial curve of a specified degree, usually linear, quadratic or cubic, to a sample window of neighboring observed values. The polynomial curve fitting process is based on a regression analysis algorithm that employs the least squares method to minimize the error values between the fitted curve and the collected data. The coefficients of the polynomial involved in the smoothing process are then used to estimate the signal value at a given point (Schmid et al., 2022; Krishnan & Seelamantula, 2012; Schafer, 2011).

In digital signal processing, the Savitzky-Golay filter is a widely used technique for data smoothing. It is specifically designed to enhance the accuracy of a data set without compromising its structural integrity (Liu et al., 2016; Azami et al., 2012; Acharya et al., 2016). The filter works by fitting a polynomial curve to a set of neighboring observations and determining a representative estimated value for the smoothed sample. To control the degree of smoothing, the window size and polynomial degree can be adjusted as parameters. However, larger window sizes or higher degree polynomials may introduce offset and overshoot, leading to inaccuracies in the estimated results.

It is crucial to consider the relationship between the window length and the degree of the polynomial filter when using a Savitzky-Golay filter. A higher polynomial degree can result in overfitting, whereas a smaller degree may not provide effective smoothing. The filter uses a polynomial of degree N to approximate M points, where M is the size of the window. Therefore, choosing N+1 to be as close as possible to M results in no actual smoothing, but instead only interpolation. It is recommended to select a value of N that is significantly smaller than M, for numerical stability and effective smoothing. Typically, a polynomial degree between 2 and 5 is used in applications employing a Savitzky-Golay filter (Quan & Cai, 2012; Awal et al., 2011; Hasan et al., 2022; Schulze et al., 2008).

The Savitzky-Golay filter is a mathematical model that utilizes a regression algorithm based on the least squares' method. It has proven to be effective in removing noises with a specific frequency spectrum and in smoothing graphical representations over long periods of time. In terms of implementation, the filter is relatively easy to define, compared to more complex filters such as the Kalman filter. As a result, it has a wide range of applications in various economic fields. The filter is known for providing minimal delay and overshoot, which is significantly superior to other smoothing techniques. This feature is particularly crucial in applications that process data in real-time or require high accuracy results.

Using the same dataset as that employed for the simple moving average filter, various combinations of window size and polynomial degree values were empirically tested to ensure that the filtering process yields accurate results, reflecting the physical displacement. It was observed that an excessively small window size prevents filtering out measurement errors, resulting in the filtered signal closely tracking the trends of the original signal. Specifically, for a window size of 11 elements, the filtered signal was nearly identical to the original one for polynomials of degree 3, 5, and 9. For window sizes around 25, the filtering process substantially improved, with measurement errors being corrected to a large extent for a polynomial of degree 3. For higher degrees, filtering accuracy decreased, with the processing technique failing to remove second-order local extrema points. The most accurate filters were obtained for a window size equal to 37 elements, according to the experiments conducted. The filtered values for both degree 3 and degree 5 polynomials closely followed the 10 steps, flattened the secondary peaks and tracked the general trend of the acceleration values. In other words, the pairs {37, 3} and {37, 5} resulted in no significant distortions in the final values and their representation. However, a high polynomial degree automatically introduces secondary extrema points in the smoothed dataset, which contradicts the purpose of mapping a displacement process, as it adds additional steps in the representation, altering the final result.

Fig. 4 illustrates the results of applying a Savitzky-Golay filter with a window size equal to 37 elements for 3 polynomials of different degrees 3, 5 and 9.

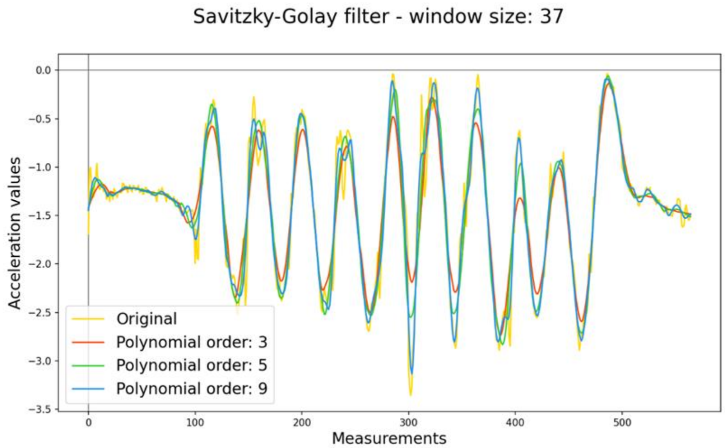

Fig. 4. Savitzky-Golay filter run on different polynomial orders

The Savitzky-Golay filter can become computationally intensive when dealing with large datasets or high-degree polynomials. The algorithm's performance heavily depends on the chosen window size and polynomial degree. Ideally, one should opt for the smallest window size and polynomial degree that yield satisfactory results, to minimize the impact on the data.

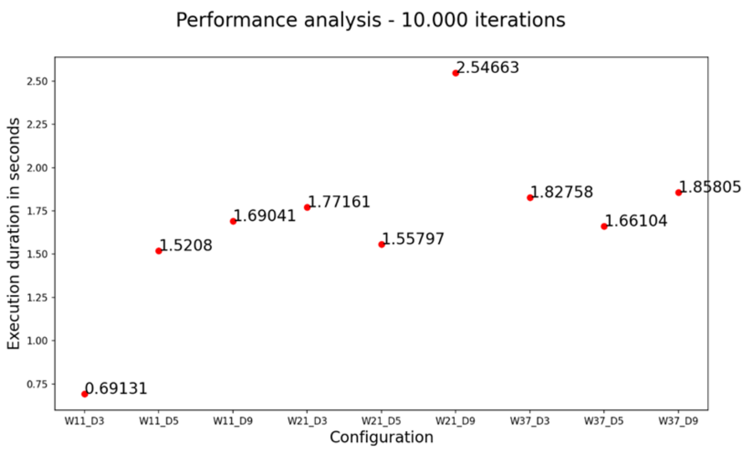

Fig. 5. Performance analysis for the Savitzky-Golay filters presented above

To evaluate the algorithm's performance, a suite of tests was conducted on the benchmark data set, filtering it with a Savitzky-Golay filter, using various configurations. Fig. 5

illustrates the results of these tests, depicting the average execution time of 10,000 simulations.

In terms of the horizontal axis, each configuration is identified by labels such as W11-D3, where W11 refers to a window size of 11 elements and D3 refers to a polynomial degree of 3. The resulting scatter plot illustrates an almost exponential increase in average execution time when the degree of the polynomial remains fixed, while the window size varies. However, when the window size remains constant and the polynomial degree varies, the average execution time exhibits a linear growth trend. It should also be mentioned that when the data set was experimentally increased to 307,360 points, the average execution time increased to an interval between 35.85593 and 111.30111 seconds, the most advantageous values being associated with a degree of the polynomial as low as possible.

Based on the results of the aforementioned tests, the pairs {37, 3} and {37, 5} were identified as potential optimal choices for filtering local extrema points. Since the objective of the filtering process is to identify extrema points in the acceleration graph that accurately reflect the actual displacement, rather than to determine the actual acceleration values, the choice of the optimal pair must also consider the average execution time. Accordingly, the pair {37, 3} is selected due to its 14-22% faster average execution time, which is crucial for a solution processing large amounts of data in real-time.

2.3. Kalman filter

If the two filters previously described are examples of FIR filters, then the third filter employed in the current analysis is an IIR implementation, as a Kalman filter simultaneously satisfies several sine quibus non-criteria. These criteria will be elaborated on in detail below.

To determine the characteristics of digital noise affecting the process of measuring acceleration values using a standard accelerometer, such as those present in smartphones, values were collected when the phone was positioned on a flat surface and no force was exerted on it, except gravity. A subset of these values is presented in Tab. 1 from the introductory chapter, and their graphical representation as a probability density function is depicted in Fig. 6.

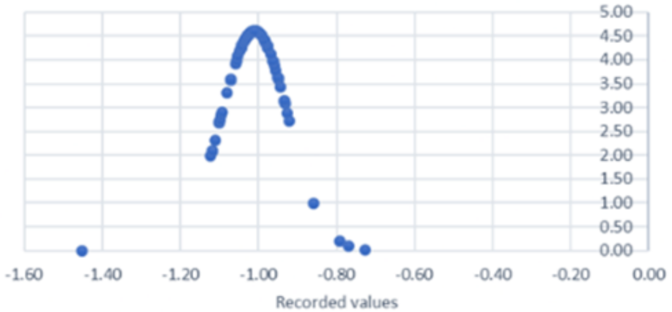

Fig. 6. Probability density function for the stationary acceleration values

It is apparent that the stationary acceleration values exhibit the characteristics of a normal distribution. The graphs previously presented also illustrate the linearity of the accelerometer, as the displacement recorded during the ten steps was a straight line. It is not feasible to make assumptions about the constancy of the acceleration values during movement or compare the values recorded by the accelerometer in the mobile phone with reference values without an additional measurement tool. Nevertheless, the amplitude highlighted graphically adheres to the cadenced movement trend, with steps of comparable lengths distributed at relatively equal time intervals. Given that sinusoids tend to repeat in a uniform and realistic manner, it is reasonable to consider the accelerometer employed in the data collection as linear.

The linear nature of the accelerometer, the Gaussian distribution of stationary values and the fact that the phenomenon being refined is time-varying, provide sufficient and necessary justifications for using an IIR implementation of a Kalman filter to filter the recorded accelerometer measurement errors.

A Kalman filter is an algorithm that uses historical observations, referred to as estimates, as well as current measurements to estimate the state of a system. These estimates and measurements are subject to error values and noise due to interfering factors. The objective of a Kalman filter is to determine the current or future state of the system (Meinhold & Singpurwalla, 1983).

The Kalman filtering process is composed of two stages. The first stage, the prediction stage, utilizes the current state estimates of the system and a mathematical prediction model to analyze the potential state of the system in the next moment of time. The prediction model represents an idealized version of the dynamic system and provides a prediction about the system's state in the next time interval, rather than an exact measurement. In the second stage, the updating stage, both the estimates obtained in the previous stage and the measurements empirically collected at the time associated with the next state are taken into consideration. The differences between the actual measurements and the estimates obtained from the model are used to determine a more precise estimate of the system state at the next time interval. This algorithm is iterative, utilizing the estimate obtained in the previous processing stage and a random estimate for the subsequent time interval (Welch, 2014).

The Kalman filter also has configurable parameters that can be adjusted depending on the specifics of the problem. If the estimation methods, used to analyze the system, are precise and provide results that are close to reality, then the Kalman gain can be set to a value as close to 1 as possible. Conversely, if the estimation methods used are imprecise, the Kalman gain will tend towards 0 (Chen et al., 2011).

Kalman filters are primarily used with time series data that describes time-varying processes. These values that characterize the process are often collected through means that induce uncertainty and measurement errors. Kalman filters are frequently used in a variety of fields, including collected data from sensors in radar, IMU or GPS applications, control systems, robotics, autonomous vehicles, financial-banking sectors, for estimating the price of shares or exchange rates, and real-time image processing, such as identifying the position and speed of a moving object in a video data stream (Li et al., 2015; Deep et al., 2018; Farhad et al., 2023; Zhang & Yu, 2022).

One of the key advantages of Kalman filters is their computational efficiency. The recursive approach used by the algorithm only considers the values determined in the previous step and new estimates, enabling it to operate efficiently within large systems, with numerous parameters. The versatility of the algorithm is demonstrated by its wide applicability across numerous economic fields. Furthermore, the system is adaptive, closely following changes that impact the state of the system over time and providing optimal estimators of the real state of a system despite the presence of noise and uncertainty (Kwon & Park, 2020).

To implement a Kalman filter, the most crucial step is to select a suitable model that describes the evolution of the analyzed system. The filter requires the choice of a mathematical model that captures the evolution and transition of the system's dynamics, as well as an appropriate measurement process. It is essential to invest sufficient time and resources in selecting the optimal variables since this stage cannot be superficially treated. The filter also assumes that the processed signals follow a Gaussian distribution, and if the collected values do not meet this requirement, a preprocessing step is mandatory. Additionally, it is often necessary to adjust and match some parameters of the model, such as the covariance of the estimation and measurement errors. Finally, it is crucial to validate the Kalman filter in relation to factual, concrete evidence to obtain accurate and reliable estimates.

In the case of the other two analyses presented earlier, the input data set underwent a filtering process, using a Kalman filter with four different configurations. The four configurations used were the result of repeated tests carried out to adjust the parameters described in the previous paragraphs. The modifications aimed to efficiently filter the input digital signal and optimize resource consumption. The changes impacted both the number of decimal places used in the definition of Δt, process noise covariance, measurement noise covariance, as well as the accuracy of the approximation. In configuration C1, the time interval Δt is 0.017, the process noise covariance matrix is a 2D matrix with values of [[0.01, 0], [0, 0.01]] and the measurement noise covariance matrix is [[1]]. In configuration C2, Δt is 0.01, the process noise covariance matrix is a 2D matrix with values of [[0.01, 0], [0, 0.01]] and the measurement noise covariance matrix is [[2]]. In configuration C3, Δt is 0.1, the process noise covariance matrix is a 2D matrix with values of [[0.1, 0], [0, 0.1]] and the measurement noise covariance matrix is [[1]]. In configuration C4, Δt is 0.017, the process noise covariance matrix is a 2D matrix with values of [[1, 0], [0, 1]] and the measurement noise covariance matrix is [[2]]. All other parameters remain constant across all configurations. The results obtained are illustrated in Fig. 7.

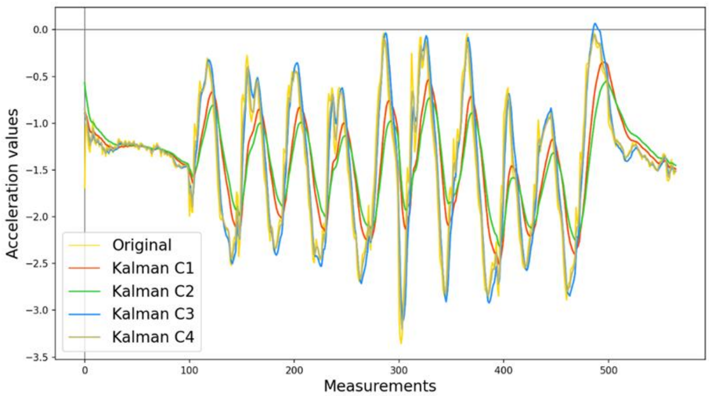

Fig. 7. Kalman filter run on different configurations

From the standpoint of filtering efficiency and performance, two out of the four configurations produced satisfactory results in detecting the recorded steps during the journey, exhibiting significant differences from the others. The outcomes of the performance tests carried out, similar to those performed for the FIR filters, are shown in Fig. 8. These tests demonstrate that configuration C1 yields the most encouraging results in terms of performance.

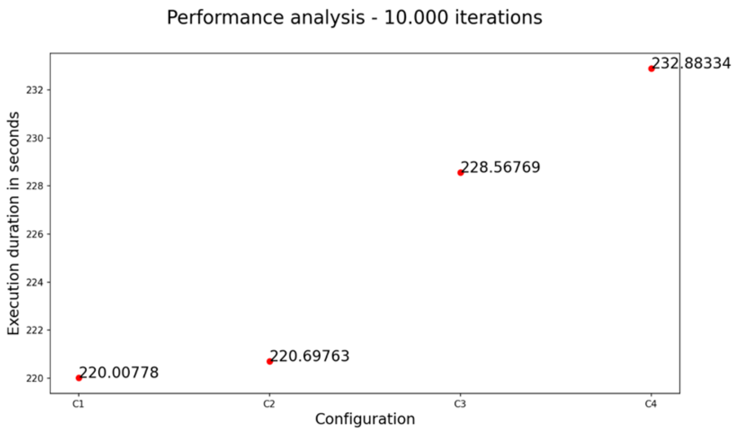

Fig. 8. Performance analysis for the Kalman filters presented above

3. RESULTS AND DISCUSSION

The findings presented in the preceding chapter may be examined and interpreted from multiple perspectives. Depending on the intended analysis objective, the current matter and how the use of a filter fits in with other components of a prospective solution, certain viewpoints may be more favorable than others and vice versa.

For instance, if the digital signal filtering stage is a constituent of a broader solution that aims to determine movement within an indoor area, using Physics theory-based means through double integration of acceleration values, then an approach that preserves acceleration values as accurately as possible, without modifying the magnitude and dynamics of the waves, is preferable to one that substantially changes the digital signal's representation. Nevertheless, such an approach is highly challenging and the filter calibration phase is crucial, necessitating extensive testing to validate the filtering process's impact and intensity, as well as the accuracy of the outcomes (Benoussaad et al., 2015).

Conversely, if the filtering stage is just one part of a more complex algorithm, designed to accurately identify moments when a target passes by, while another component that uses different techniques determines the distance without relying on double integration of acceleration, filtering techniques that minimize noise and errors are more appealing in achieving this objective (Dhanalakshmi et al., 2023).

The choice of which technique to use must simultaneously consider several factors. The most important of these factors is the nature and type of digital signal being processed. Specific characteristics such as amplitude, frequency, collection method and error intensity must be considered when deciding which filter to use. The filtering approach is also critical. Is the filter applied in real-time to a continuous data stream or is it retrospectively applied to past data? What is the purpose of using a filter: is it integrated into an application that needs to provide a real-time response or is the response's accuracy and correctness more critical than speed? What is the order of magnitude of the dataset being filtered: is it hundreds of observations or a large volume of hundreds of thousands or millions of observations? What hardware resources are available for a solution that will incorporate the filter?

The incorporation of sensor data processing methodologies into Internet of Things applications is crucial for improving both the reliability and precision of these systems. These systems are progressively being utilized in essential domains, including health monitoring, smart urban development, and industrial automation (Al-Fuqaha et al., 2015; Ray, 2018). The intrinsic variability present in IoT settings, marked by an array of sensor types, deployment magnitudes, and application-specific demands, requires an adaptable strategy towards data filtering. For example, within the context of health monitoring systems, the simplicity and reduced computational requirements of the Simple Moving Average (SMA) technique render it a favorable option for systems dedicated to real-time patient surveillance, where there are typically constraints on power usage and processing capacity (Majumder et al., 2017).

Nonetheless, the tendency of the Simple Moving Average (SMA) to excessively smooth data might render it inadequate for applications that demand precise signal analysis, like gait analysis or fall detection. In these instances, the use of Savitzky-Golay filters could offer a more optimal equilibrium between data smoothing and the preservation of signal characteristics (Li et al., 2020).

Conversely, Kalman filters, characterized by their greater computational complexity, are more apt for scenarios where precision and the capacity to forecast future states through a model of the system's dynamics hold supreme importance. Their application in navigation and tracking systems within smart urban environments underscores the crucial balance between computational expenditure and filtering efficacy, particularly when managing extensive datasets and the necessities of real-time processing (Oteafy & Hassanein, 2018).

The conversation regarding the optimization of filtering methods for particular IoT applications should also consider the scalability of these approaches. With the expansion of IoT devices, leading to the production of immense quantities of data, the computational and energy expenses linked to data processing grow more consequential (Zanella et al., 2014). Methods such as the Savitzky-Golay filter, which allow for the adjustment of window size and polynomial degree, present a way to equilibrate the computational burden with the effectiveness of filtering, an essential factor in deployments on a large scale.

Tab. 2. Relevant performance indicators of presented methods

C1	C2	C3	C4	C5	C6
SMA W7	L	L	M	NO	0.06019
SMA W13	L	M	H	YES	0.06325
SMA W37	L	VH	M	NO	0.08229
SavGol W37 D3	M	H	H	YES	1.82758
SavGol W37 D5	M	M	H	YES	1.66104
SavGol W37 D9	M	L	M	NO	1.85805
Kalman C1	H	H	H	YES	220.00
Kalman C2	H	H	H	YES	220.69
Kalman C3	H	L	M	NO	228.56
Kalman C4	H	L	M	NO	232.88

Tab. 2 summarizes the measured values from the Materials and Methods section, where column names represent: C1 - Filter name, C2 - Implementation complexity, C3 - Impact on original signal, C4 - Capacity to effectively reduce noise, C5 - Correctly identified the number of steps and C6 - 10000 iterations duration in seconds.

4. CONCLUSIONS

By means of a comparative analysis of the tabular values from Tab. 2, it is evident that for an application that seeks solely to accurately identify the number of steps taken during a journey, with separate time intervals for populations of 500 observations, one may choose between a simple moving average (SMA) filter with a window size of 13 entities, a Savitzky-Golay (SavGol) filter with a window size of 37 and a polynomial degree of 3 or 5, or a Kalman filter with a configuration similar to C1 or C2. If real-time processing that does not significantly distort the signal is desired, then the number of viable options is reduced to 2: SMA with window size 13 and SavGol W37 D5.

In contrast, if the intention is to perform double integration of the acceleration, the options to consider are: SMA W7, SavGol W37 D9, Kalman C3 and Kalman C4. However, depending on the size of the sample being processed, certain implementations may become unsatisfactory. It is important to note that there is no single technique that is universally applicable to all possible scenarios.

In the context of this work, where the filtered signal was obtained using an accelerometer available in a smart mobile phone, that presents significant measurement errors when collecting data in a stationary frame, with a mean square deviation of 0.687934, achieving the double integration of the acceleration values is difficult. Instead, a more feasible approach is to use models to estimate the average step and stride length based on the subject's height, gender and weight. Furthermore, since real-time monitoring of movement is necessary, a Savitzky-Golay filter with a window size of 37 and a polynomial degree of 5 would be a viable option for this specific scenario.

Such a module for collecting data from a sensor and filtering digital signals to identify the steps taken can be integrated into future developments of an application designed for indoor localization in large buildings with frequent points of interest. The application can monitor the subject's position in real-time and serve as a personal assistant that provides suggestions and contextual recommendations tailored to the situation.

Acknowledgments

This paper was co-financed by The Bucharest University of Economic Studies during the PhD program.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

REFERENCES

Acharya, D., Rani, A., Agarwal, S., & Singh, V. (2016). Application of adaptive Savitzky–Golay filter for EEG signal processing. *Perspectives in science*, *8*, 677-679. https://doi.org/10.1016/j.pisc.2016.06.056

Ahlawat, J., Sharma, M., & Pundir, C. S. (2024). Advances in xanthine biosensors and sensors: A review. *Enzyme and microbial technology*, *174*, 110377. https://doi.org/10.1016/j.enzmictec.2023.110377

Ahn, H. S., & Ko, K. H. (2009). Simple pedestrian localization algorithms based on distributed wireless sensor networks. *IEEE Transactions on Industrial Electronics*, *56*, 4296-4302. https://doi.org/10.1109/TIE.2009.2017097

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, *17*(4), 2347-2376. https://doi.org/10.1109/COMST.2015.2444095

Anwar, A., Anwar, A., Moukahal, L., & Zulkernine, M. (2023). Security assessment of in-vehicle communication protocols. *Vehicular Communications*, *44*, 100639. https://doi.org/10.1016/j.vehcom.2023.100639

Awal, M. A., Mostafa, S. S., & Ahmad, M. (2011). Performance analysis of Savitzky-Golay smoothing filter using ECG signal. *International Journal of Computer and Information Technology*, *1*(02), 110126.

Azam, S., Dall'Ora, N., Fraccaroli, E., Gillon, R., & Fummi, F. (2023). Analog defect injection and fault simulation techniques: A systematic literature review. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *43*(1), 16-29. https://doi.org/10.1109/TCAD.2023.3298698

Azami, H., Mohammadi, K., & Bozorgtabar, B. (2012). An improved signal segmentation using moving average and Savitzky-Golay filter. *Journal of Signal and Information Processing*, *03*, 39-44. https://doi.org/10.4236/JSIP.2012.31006

Benoussaad, M., Sijobert, B., Mombaur, K. D., & Azevedo Coste, C. (2015). Robust foot clearance estimation based on the integration of Foot-Mounted IMU acceleration data. *Sensors*, *16*(1), 12. https://doi.org/10.3390/s16010012

Bentler, R., & Chiou, L. K. (2006). Digital noise reduction: An overview. *Trends in amplification*, *10*(2), 67-82. https://doi.org/10.1177/1084713806289514

Bloecher, H. L., Dickmann, J., & Andres, M. (2009). Automotive active safety & comfort functions using radar. *2009 IEEE International Conference on Ultra-Wideband* (pp. 490-494). IEEE. https://doi.org/10.1109/ICUWB.2009.5288790

Chen, Z., Sun, Z., & Wang, W. (2011). Design and implementation of Kalman filter. *IET International Conference on Communication Technology and Application (ICCTA 2011)* (pp. 901-904). https://doi.org/10.1049/cp.2011.0800

Deep, A., Mittal, M., & Mittal, V. (2018). Application of Kalman filter in GPS position estimation. *2018 IEEE 8th Power India International Conference (PIICON)* (pp. 1-5). IEEE. https://doi.org/10.1109/POWERI.2018.8704368

Dhanalakshmi, S. B., Joseph, C., Rajeswari, A., Poongothai, M., & Rozario, J. (2023). Enhancing indoor navigation: bluetooth Beacon-Based localization systems. *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 133-138). IEEE. https://doi.org/10.1109/ICECA58529.2023.10395279

Dinakar, J. R., & Vagdevi, S. (2017). A study on storage mechanism for heterogeneous sensor data on big data paradigm. *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 342-345). IEEE. https:/doi.org/10.1109/ICEECCOT.2017.8284525

Ellis, C. A., & Parbery, S. A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance*, *19*(3), 399-411. https://doi.org/10.1016/j.ribaf.2004.12.009

Farhad, A., Kwon, G. R., & Pyun, J. Y. (2023). Mobility adaptive data rate based on Kalman filter for LoRa-Empowered IoT applications. *2023 IEEE 20th Consumer Communications & Networking Conference (CCNC)* (pp. 321-324). IEEE. https://doi.org/10.1109/CCNC51644.2023.10060330

Gujarathi, T., & Bhole, K. (2019). Gait analysis using imu sensor. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE. https://doi.org/10.1109/ICCCNT45670.2019.8944545

Hansun, S., & Kristanda, M. B. (2017). Performance analysis of conventional moving average methods in forex forecasting. *2017 International Conference on Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS)* (pp. 11-17). IEEE. https://doi.org/10.1109/ICON-SONICS.2017.8267814

Hasan, K., Meraj, S. T., Othman, M. M., Lipu, M. H., Hannan, M. A., & Muttaqi, K. M. (2022). Savitzky–Golay Filter-Based PLL: Modeling and performance validation. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1-6. https://doi.org/10.1109/TIM.2022.3196946

Hidayat, A. A., Arief, Z., & Happyanto, D. C. (2015). Mobile application with simple moving average filtering for monitoring finger muscles therapy of post-stroke people. *2015 International Electronics Symposium (IES)* (pp. 1-6). IEEE. https://doi.org/10.1109/ELECSYM.2015.7380803

Johnston, F. R., Boyland, J. E., Meadows, M., & Shale, E. (1999). Some properties of a simple moving average when applied to forecasting a time series. *Journal of the Operational Research Society*, *50*(12), 1267-1271. https://doi.org/10.2307/3010636

Karpinski, R., Krakowski, P., Jonak, J., Machirowska, A., & Maciejewski, M. (2023). Comparison of selected classification methods based on machine learning as a diagonostic tool for knee joint cartilage damage based on generated vibroacoustic processes. Applied Computer Science, *19*(4), 136–150. https://doi.org/10.35784/acs-2023-40

Krishnan, S. R., & Seelamantula, C. S. (2012). On the selection of optimum Savitzky-Golay filters. *IEEE transactions on signal processing*, *61*(2), 380-391. https://doi.org/10.1109/TSP.2012.2225055

Kwon, J., & Park, D. (2020). Implementation of computation-efficient sensor network for Kalman filter-based intelligent position-aware application. *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)* (pp. 565-568). IEEE. https://doi.org/10.1109/ICAIIC48513.2020.9065098

Li, Q., Li, R., Ji, K., & Dai, W. (2015). Kalman filter and its application. *2015 8th International Conference on Intelligent Networks and Intelligent Systems (ICINIS)* (pp. 74-77). IEEE. https://doi.org/10.1109/ICINIS.2015.35

Li, Q., Rajagopalan, C., & Clifford, G. D. (2020). Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Transactions on Biomedical Engineering*, *61*(6), 1607-1613. https://doi.org/10.1109/TBME.2013.2275000

Liu, Y., Dang, B., Li, Y., Lin, H., & Ma, H. (2016). Applications of Savitzky-Golay filter for seismic random noise reduction. *Acta Geophysica, 64,* 101-124. https://doi.org/10.1515/acgeo-2015-0062

Majumder, S., Aghayi, E., Noferesti, M., Memarzadeh-Tehran, H., Mondal, T., Pang, Z., & Deen, M. J. (2017). Smart homes for elderly healthcare—Recent advances and research challenges. *Sensors, 17*(11), 2496. https://doi.org/10.3390/s17112496

Meinhold, R. J., & Singpurwalla, N. D. (1983). Understanding the Kalman filter. *The American Statistician*, *37*(2), 123-127. https://doi.org/10.2307/2685871

Mohamed, A. M., Fouad, F. H., Raouf Fayek, G., El Sayed, K. M., Ahmed, M. N., Mahmoud, R. Z., & El Nashar, R. M. (2023). Recent advances in electrochemical sensors based on nanomaterials for detection of red dyes in food products: A review. *Food chemistry, 435*, 137656. https://doi.org/10.1016/j.foodchem.2023.137656

Olivares, A., Olivares, G., Gorriz, J. M., & Ramirez, J. (2009). High-efficiency low-cost accelerometer-aided gyroscope calibration. *2009 International Conference on Test and Measurement* (pp. 354-360). IEEE. https://doi.org/10.1109/ICTM.2009.5412920

Oteafy, S. M. A., & Hassanein, H. S. (2018). Leveraging Tactile Internet Cognizance and Operation via IoT and Edge Technologies. *IEEE*, *107*(2), 364-375. https://doi.org/10.1109/JPROC.2018.2873577

Poulose, A., Kim, J., & Han, D. S. (2019). Indoor localization with smartphones: Magnetometer calibration. *2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-3). IEEE. https://doi.org/10.1109/ICCE.2019.8661986

Promrit, P., Chokchaitam, S., & Ikura, M. (2018). In-vehicle MEMS IMU calibration using accelerometer. *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)* (pp. 1-3). IEEE. https://doi.org/10.1109/ICSIMA.2018.8688778

Purnama, S. I., Afandi, M. A., & Purba, E. V. (2022). Global Positioning System Data Processing Improvement for Blind Tracker Device Based Using Moving Average Filter. In T. Triwiyanto, A. Rizal, & W. Caesarendra (Eds.), *Proceedings of the 2nd International Conference on Electronics, Biomedical Engineering, and Health Informatics* (pp. 177–188). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-1804-9_14

Quan, Q., & Cai, K. Y. (2012). Time-domain analysis of the Savitzky–Golay filters. *Digital Signal Processing*, *22*(2), 238-245. https://doi.org/10.1016/j.dsp.2011.11.004

Ray, P. P. (2018). A survey on Internet of Things architectures. *Journal of King Saud University - Computer and Information Sciences, 30*(3), 291-319. https://doi.org/10.1016/j.jksuci.2016.10.003

Schafer, R. W. (2011). On the frequency-domain properties of Savitzky-Golay filters. *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)* (pp. 54-59). IEEE. https://doi.org/10.1109/DSP-SPE.2011.5739186

Schmid, M., Rath, D., & Diebold, U. (2022). Why and how Savitzky–Golay filters should be replaced. *ACS Measurement Science Au, 2*(2), 185-196. https://doi.org/10.1021/acsmeasuresciau.1c00054

Schulze, H. G., Foist, R. B., Ivanov, A., & Turner, R. F. (2008). Fully automated high-performance signal-to-noise ratio enhancement based on an iterative three-point zero-order Savitzky–Golay filter. *Applied spectroscopy*, *62*(10), 1160-1166. https://doi.org/10.1366/000370208786049079

Sung, F. Y., Fang, S. H., & Chien, Y. R. (2014). An experimental study of MEMS-based magnetometers on Android mobile phones. *2014 IEEE International Conference on Consumer Electronics – Taiwan* (pp. 227-228). IEEE. https://doi.org/10.1109/ICCE-TW.2014.6904071

Tan, L., & Jiang, J. (2013). *Digital signal processing: Fundamentals and applications*. Academic Press.

Thinh, D. T., Quan, N. B. H., & Maneetien, N. (2018). Implementation of moving average filter on STM32F4 for vibration sensor application. *2018 4th International Conference on Green Technology and Sustainable Development (GTSD)* (pp. 627-631). IEEE. https://doi.org/10.1109/GTSD.2018.8595630

Truzman, S., Revach, G., & Klein, I. (2021). On the influence of home appliances on the smartphone's inertial sensors. *2021 IEEE Sensors*, 1-4. IEEE. https:/doi.org/10.1109/SENSORS47087.2021.9639739

Vaseghi, S. V. (2008). *Advanced digital signal processing and noise reduction*. John Wiley & Sons.

Welch, G. F. (2014). Kalman Filter. In K. Ikeuchi (Ed.), *Computer Vision*. Springer US. https://doi.org/10.1007/978-0-387-31439-6_716

Xie, Q., Wang, Q., Cao, N., Gao, S., Liang, G., Zhang, T., Chen, Y., Zhai, Y., & Li, N. (2017). A survey of wireless sensor technique applications for medical mare. *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)* (pp. 412-415). IEEE. https://doi.org/10.1109/CyberC.2017.87

Yadav, L., Yadav, P., Laddha, H., Sharma, M., Sharma, P., Agarwal, M., & Gupta, R. (2024). Smartphone based CN– ion sensing in electroplating wastewater and food samples in a simple, economical and rapid manner. *Inorganic Chemistry Communications*, *161*, 112083. https://doi.org/doi:10.1016/j.inoche.2024.112083

Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of Things for smart cities. *IEEE Internet of Things Journal, 1*(1), 22-32. https://doi.org/10.1109/JIOT.2014.2306328

Zhang, P., Gu, J., Milios, E. E., & Huynh, P. (2005). Navigation with IMU/GPS/digital compass with unscented Kalman filter. *IEEE International Conference Mechatronics and Automation, 2005* (pp. 1497-1502). IEEE. https://doi.org/10.1109/ICMA.2005.1626777

Zhang, X., & Yu, W. (2022). Research on the application of Kalman Filter algorithm in aircraft trajectory analysis. *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)* (pp. 196-199). IEEE. https://doi.org/10.1109/ICSP54964.2022.9778746