



HAL
open science

D-SORM: A digital solution for remote monitoring based on the attitude of wearable devices

Manuel Abbas, Dominique Somme, Regine Le Bouquin Jeannes

► **To cite this version:**

Manuel Abbas, Dominique Somme, Regine Le Bouquin Jeannes. D-SORM: A digital solution for remote monitoring based on the attitude of wearable devices. *Computer Methods and Programs in Biomedicine*, 2021, 208, 10.1016/j.cmpb.2021.106247 . hal-03329645

HAL Id: hal-03329645

<https://hal.science/hal-03329645v1>

Submitted on 7 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Highlights

- Digital clinical solution to help medical teams
- Device attitude and 3D visualization
- Arm tele-rehabilitation using wearable technologies
- Human motion tracking for elderly care
- Novel machine learning architecture using feature fusion

Journal Pre-proof

D-SORM: A Digital Solution for Remote Monitoring Based on the Attitude of Wearable Devices

Manuel Abbas^{a,*}, Dominique Somme^b and Régine Le Bouquin Jeannès^a

^aLTSI-UMR 1099, Inserm, Université de Rennes 1, 35000 Rennes, France

^bCHU de Rennes, Service de Gériatrie, Rennes, 35000, France

ARTICLE INFO

Keywords:

E-health
Device attitude
Sensor fusion
Arm tele-rehabilitation
Activity recognition
Machine learning

ABSTRACT

Background and objective: E-health is a growing research topic, especially with the expansion of the Internet of Things (IoT). Miniaturized wearable sensors are auspicious tools for biomedicine and healthcare systems. In this paper, we present D-SORM, a sensor fusion-based digital solution intended to assist clinicians and improve their diagnosis by providing objective measurements and automatic recognition. The aim is to supply an interface for remote monitoring to the medical staff.

Methods: D-SORM platform estimates the wearable device attitude based on its acquired data, and visualizes it in real-time using a graphical user interface (GUI). It also integrates two modules which serve two different medical applications. The first one is arm tele-rehabilitation, where sessions are done online. The practitioner gives the instructions while wearing the device, and the patient has to reproduce the gestures. A processing unit is dedicated to compute statistical features and calculate the success rate. The second one is human motion tracking for elderly care. A novel machine learning architecture is proposed, based on feature fusion, to predict the activities of daily living.

Results: The rehabilitation mechanism was tested under supervised conditions, by performing a set of movements. D-SORM provides extra information and objective measurements, thus facilitates the diagnosis of clinicians. The human activity recognition is also validated using a public dataset. With D-SORM, an efficiency ranging from 97.7% to 99.65% is ensured under unsupervised conditions.


Conclusion: The proposed design constitutes a digital clinical tool for medical teams allowing remote health monitoring. It overcomes geographical barriers while providing faster and highly accurate assessment.

1. Introduction

E-health is the use of communication technologies for healthcare [33]. It ensures remote delivery of wellness programs and services, using information and communication technologies for diagnosis, treatment, and prevention of disease and injuries. The healthcare systems are shifting towards this practice of caring, especially during the covid-19 pandemic which we are currently facing, since the provider and the patient are not physically present in the same place. This process has been

used in the past [27] and is currently a trending research topic to overcome geographical barriers, especially with the occurrence of the population aging phenomenon. In the past two decades, a wide range of medical applications have been developed using Internet of Things (IoT) and wearable technologies [35, 14]. The vital signs monitoring approaches are broadly used in the medical context. Measurements of body temperature [26], heart rate [34], blood pressure [5], and respiration rate [8] are critical for people suffering from chronic diseases. Indoor positioning [31, 13] is also a priority for health management, particularly for elderly care. In this case, older adults suffering from dementia, or vulnerable subjects who regularly fall, are found in a timely manner. These techniques are based on technologies such as Infrared Ray, RFID, Wifi/WLAN,

*Corresponding author

 <https://orcid.org/0000-0003-0939-0781> (M. Abbas);

dominique.somme@chu-rennes.fr (D. Somme);

regine.le-bouquin-jeannes@univ-rennes1.fr (R. Le Bouquin Jeannès)

ORCID(s): 0000-0003-0939-0781 (M. Abbas);

0000-0002-4050-2895 (R. Le Bouquin Jeannès)

to name a few. More recently, the surveillance of human motion and gesture recognition via wearable sensors, have emerged as robust tools for a wide-range of healthcare applications [1, 4].

This present work proposes a new digital solution to assist and improve the diagnosis of clinicians by visualizing the human's movements and providing objective measurements, using a user-friendly interface. This paper describes the workflow of the system, and presents the experimental frame to validate the efficiency of the processing units.

1.1. Related Work

To monitor physiological signals, an inertial system is needed, which consists of one or more sensors. The widely used sensing devices are 3D accelerometers, measuring the acceleration, 3D gyroscopes, measuring the angular velocity, and 3D magnetometers, measuring magnetic fields. By fusing data from the aforementioned sensors, informative features could be computed to recognize and monitor human gestures. This constitutes the basis of different assistive healthcare solutions, like tele-rehabilitation (TR), fall detection, and activity monitoring, to deliver therapy and/or recommendations. Jamwal *et al.* [11] developed a wearable ankle robot to ensure remote rehabilitation programs, with a library of exercises creating virtual reality. BrÅijckner *et al.* [3] presented an onboard inertial sensor fusion approach for home based stroke rehabilitation. A wearable platform was proposed, featuring an interface and a sound synthesis toolkit linked to wireless hearing aids. Pierleoni *et al.* [25] tackled fall detection by forming an attitude heading reference system (AHRS). Alarms generated by their algorithm are sent via the Bluetooth module, and a mobile application was developed to make pre-recorded phone calls after receiving alarms from the sensor node. Furthermore, He *et al.* [10] created an alerting system which distinguishes falls from activities of daily living (ADLs). As soon as the system detects a fall, an alarm is sent to caregivers to provide timely and accurate help for the elderly. Human activity recognition (HAR) [23, 2] is another application which serves health purposes. It has been widely discussed in the past two decades, and is still an active research area [36]. Tracking ADLs continuously allows the detection of abnormal events, progressive weakening in elderly, and

constitutes the basis of prevention systems (like fall and/or frailty prevention). Besides, the recognition of physical activities provides daily reports consisting of important metrics like the number of steps and the movement intensity [20].

Even though health informatics has been a trending research topic, interfaces serving the medical staff during their examination have received little attention. Moreover, there are still some concerns which need to be improved and fully addressed. Detecting and localizing falls or abnormal events are important to provide immediate help to the older person. Nevertheless, the clinician must be able to observe the succession of ADLs which led to this abnormal event in order to assess the health status of the subject and to deliver needed recommendations. Furthermore, he might need to visualize the human motion for an efficient assessment. Additionally, simplifying the user interface and interactions should also be targeted.

1.2. Workflow and Main Contributions

D-SORM (Digital SOLUTION for Remote Monitoring) is a platform that is based on the orientation estimation of a wearable inertial system and could be used in a clinical environment. It ensures a smooth visualization of the human movements in a 3D space. D-SORM operates under supervised and unsupervised conditions, while featuring two modes, namely live and replay mode. In other words, the orientation can be estimated by processing (a) acquired data in real-time (live mode) while being connected to the wearable device, or (b) stored data which were already acquired prior to a certain event. The device attitude can be exploited following two different ways to serve two medical applications. The first module is arm tele-rehabilitation (TR). With one or more devices placed on the patient's arm, this unit computes statistical features to quantify the similarity between the simulated gestures and the ground truth (GT), in order to evaluate the progress of the patient and provide some feedback. The second module is activity monitoring for elderly care. The performed ADLs are recognized accurately, using one wearable device placed on the trunk of the human body, thanks to a novel and highly efficient machine learning architecture, which outperforms state-of-the-art methods. All these instructions and commands are

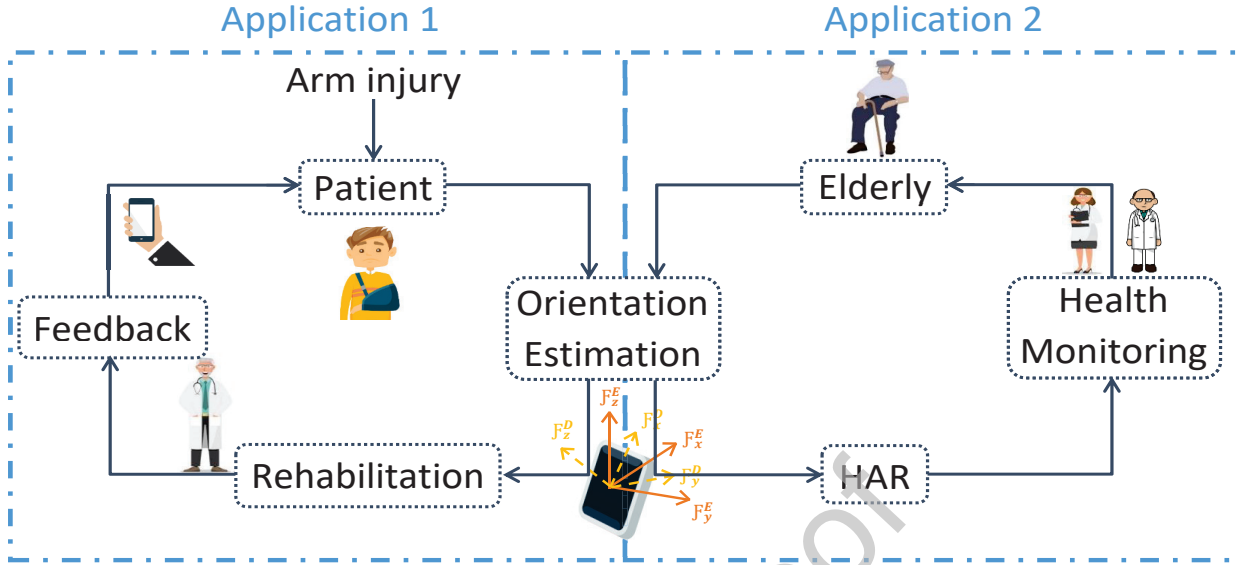


Figure 1: A flow diagram illustrating the applications of the proposed approach, as well as the workflow of D-SORM.

manipulated using a graphical user interface (GUI) which is very simple to handle. It is worth mentioning that all the aforementioned operations, constituting the process of this solution, are done without revealing the identity of the patient/elderly respecting his privacy. Note that arm TR is needed for injury and/or illness treatment, caused by a stroke/fall for example, while HAR serves elderly care like frailty/fall prevention, by tracking the physical function of the person. Hence, D-SORM considers both curative and preventive approaches. Figure 1 illustrates the workflow of this solution, and the need for such approach in the medical context.

The main contributions of this work are fourfold:

- Proposal of a sensor fusion-based algorithm to visualize the sensing unit in a 3D space.
- Proposal of a low-cost technique to analyze the performed gestures for an arm TR module.
- Proposal of a novel HAR system for health monitoring and elderly care.
- Development of a user-friendly GUI,

featuring live and replay modes, under (un)supervised conditions.

The remainder of the paper is organized as follows. Section II recalls the state-of-the-art in attitude estimation and describes some important operations needed for this task. Section III presents the sensor fusion technique and 3D visualization process. The proposed techniques regarding the two telemedicine applications are explained in details in section IV. The experimental results are shown in section V before concluding the paper in section VI.

2. Orientation Estimation: State-of-the-Art

2.1. Euler angles and Gimbal lock

To represent the 3D orientation of a body in regard to a fixed coordinate system, the Euler angles are generally considered. Those angles are denoted as ψ , θ , and ϕ . In the literature, the notion of gimbal lock has been introduced in mathematics to measure these angles. A gimbal is a ring which rotates around an axis. Hence, three gimbals are nested

one within another to ensure 3D rotations. Therefore, the orientation is described as three axial rotations with Euler angles, *i.e.* applying a fixed set of successive rotations (each rotation around a specific axis). This operation can cause an issue called Gimbal lock. This phenomenon occurs in certain situations, where two gimbals line up (the axes of two gimbals become parallel). This leads to the loss of one degree of freedom, and additional rotations from this point might lead to unexpected outcomes. A solution to this problem is the use of quaternions representation as described below.

2.2. Quaternions

Quaternions are a 4-dimensional extension of the complex numbers [15]. They can represent the orientation of a body in three-dimensional space, and produce a 3D rotation. A quaternion q is equal to $a + bi + cj + dk$, where i, j, k are fundamental quaternion units. a is called the real part and the rest ($bi + cj + dk$) is called the vector part. The conjugate \bar{q} is equal to q after multiplying the vector part by -1 . The multiplication rules regarding

this representation are as follows:

$$\begin{cases} i^2 = j^2 = k^2 = -1 \\ ij = -ji = k \\ jk = -kj = i \\ ki = -ik = j \end{cases}$$

Two fundamental units being multiplied together do not commute, since one is the negation of the other.

To describe an orientation, the quaternion should be of unit length: $q_o = \cos(\alpha) + \sin(\alpha)(r_x i + r_y j + r_z k)$, assuming that $r_x^2 + r_y^2 + r_z^2 = 1$. Three observations can be noted:

- by left-multiplying every point in the space by q_o , we rotate two circles, namely (a) the one passing by 1 and $r_x i + r_y j + r_z k$, and (b) the one perpendicular to it (these two circles are orthogonal).
- by left-multiplying every point by \bar{q}_o , we rotate these two circles in opposite directions.
- by right-multiplying by \bar{q}_o , we only negate the rotation direction of circle (a).

Hence, left-multiplying by q_o and right-multiplying by \bar{q}_o will cancel rotation (a) while doubling rotation (b) (rotation by 2α).

Now, let us project these representations to our actual situation. Figure 2 illustrates a body (like the wearable device) in a three-dimensional space. F^D denotes the device frame while F^E that of the earth. Suppose that a certain point of this body is represented by a three dimensional vector ϑ . Moreover, a certain orientation ${}^D_E q$ of a frame E relative to frame D is done through a rotation of angle θ around an axis \tilde{d} defined in D :

$${}^D_E q = [q_1 \quad q_2 \quad q_3 \quad q_4]$$

ϑ_D and ϑ_E are the description of the vector ϑ as a pure quaternion in frame D and frame E respectively, *i.e.* a value of 0 is inserted as a first element to make it a quaternion. Based on the previous remarks, the 3D rotation is ensured using the following equation:

$$\vartheta_E = {}^D_E q \otimes \vartheta_D \otimes {}^D_E \bar{q} \quad (1)$$

with: ${}^D_E q = [\cos \frac{\theta}{2} \quad -\tilde{d}_x \sin \frac{\theta}{2} \quad -\tilde{d}_y \sin \frac{\theta}{2} \quad -\tilde{d}_z \sin \frac{\theta}{2}]$ and \otimes the quaternion product.

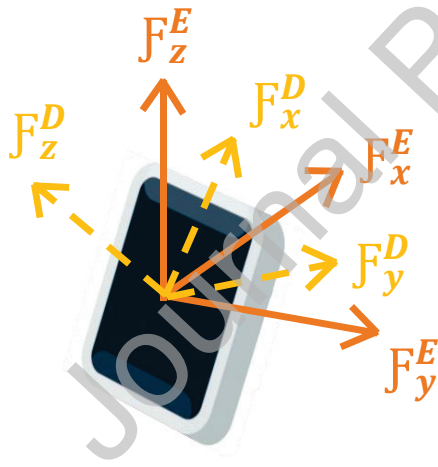


Figure 2: The sensing device with its corresponding frame (D as right superscript) and the Earth frame (E as right superscript).

Consequently, Euler angles describe this orientation achieved by sequential rotations of ψ around z_E , θ around y_E , and ϕ around x_E [15, 16]:

$$\begin{cases} \psi = \text{Atan2}(2q_2q_3 - 2q_1q_4, 2q_1^2 + 2q_2^2 - 1) \\ \theta = -\sin^{-1}(2q_2q_4 + 2q_1q_3) \\ \phi = \text{Atan2}(2q_3q_4 - 2q_1q_2, 2q_1^2 + 2q_4^2 - 1) \end{cases} \quad (2)$$

Accordingly, quaternions can solve the Gimbal lock issue. Applying quaternions can be seen as a direct transformation from a certain orientation to another one, rather than breaking this transformation into a series of rotations.

2.3. Orientation Filter

Accelerometers and gyroscopes, constituting an inertial measurement unit (IMU), are able to track translational and rotational movements. Hence, this unit measures the attitude relative to the direction of gravity. A MARG (Magnetic, Angular Rate, Gravity) sensor is the combination of an IMU and a tri-axial magnetometer. Here, a complete measurement of orientation is provided, relative to the direction of gravity and the earth magnetic field. Consequently, we consider MARG sensors in our study.

In the literature, the Kalman filter became the cornerstone for the greater part of orientation systems [7, 19, 28]. Now, despite their high accuracy, the complex implementation and high computational complexity of such systems are major drawbacks. To operate in real-time, a very low latency is required for orientation estimation approaches. Besides, other approaches have been proposed to estimate the orientation. Mahony *et al.* [18] proposed an algorithm which employs a nonlinear complementary filter. Fourati *et al.* [6] combined a quaternion-based nonlinear filter with the Levenberg Marquardt Algorithm. The algebraic quaternion algorithm (AQUA) has been also discussed [32]. In 2011, Madgwick *et al.* proposed a computationally efficient MARG orientation filter [17], based on an optimised gradient descent algorithm (GDA). It exploits quaternion representation. A detailed explanation of this algorithm is out of the scope of this paper, but we can sum it up in three steps. In the first step, data is read from the MARG sensor. Then, the GDA is applied to compute orientation increment (gradient step) using acceleration data, and magnetometer data for magnetic dis-

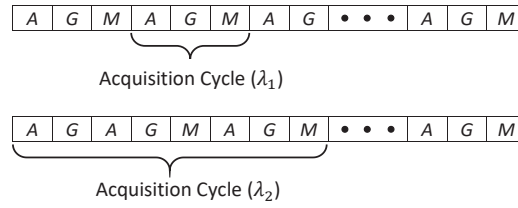


Figure 3: The acquisition cycle when (i) all modalities have the same sampling rate (λ_1) and (ii) the magnetometer has a lower rate (λ_2). In this example, ${}^A F_s = {}^G F_s = 60$ Hz and ${}^M F_s = 40$ Hz (for λ_2).

tortion compensation. The direction of the gyroscope measurement error is then computed using numerical integration. The two aforementioned operations constitute the second step. Both measurements (from the second step) are fused in a third step to estimate the attitude. Note that, conventionally, the direction of gravity \hat{g}_E defines the vertical z-axis (eq. (3)), and the earth magnetic field \hat{b}_E is considered to have components following the horizontal x-axis and the vertical z-axis (eq. (4)) [16]:

$$\hat{g}_E = [0 \quad 0 \quad 0 \quad 1] \quad (3)$$

$$\hat{b}_E = [0 \quad b_x \quad 0 \quad b_z] \quad (4)$$

3. Sensor Fusion and Real-time 3D Visualization Algorithm

A MARG wearable device, consisting of three sensors, is needed for the upcoming tasks. Those sensors are (a) a tri-axial MEMS accelerometer with a sampling frequency ${}^A F_s$ and a measurement range of ± 8 g, (b) a tri-axial gyroscope with a sampling rate ${}^G F_s$ and an angular rate of ± 2000 dps, and (c) a tri-axial magnetometer with a sampling rate ${}^M F_s$ and a full scale magnetic field of ± 4 Gauss. Two configurations (λ_i) are considered in this paper:

$$\begin{cases} {}^A F_s = {}^G F_s = {}^M F_s & (\lambda_1) \\ {}^A F_s = {}^G F_s > {}^M F_s & (\lambda_2) \end{cases}$$

Since magnetometers measure magnetic fields, their sampling frequency could be relatively low compared to the other two sensors which measure the acceleration and the angular velocity (translational and rotational movements). It is recommended to

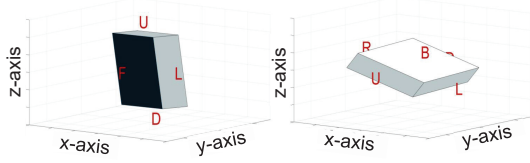


Figure 4: The orientation of the wearable sensor when the subject is walking (left graph), when the human body is hitting the ground during the impact phase of the fall (right graph).

use the same sampling rate for the latter sensors. Note that the targeted orientation filter achieves similar levels of performance from a rate of 50 Hz [16]. Figure 3 illustrates the order of data-points acquisition (acquisition cycle) following both configurations.

To display the wearable device in real-time, a platform was developed. (x, z) is the plane of the screen, and y is the axis perpendicular to the aforementioned plane. The enclosure is symbolized by a parallelepiped centered at the origin, by defining the coordinates of 8 vertices $V_i(x_i; y_i; z_i)$, and connecting them using the map M :

$$V_i = \begin{cases} V_1(-1; -0.25; -0.75) \\ V_2(1; -0.25; -0.75) \\ V_3(1; 0.25; -0.75) \\ V_4(-1; 0.25; -0.75) \\ V_5(-1; -0.25; 0.75) \\ V_6(1; -0.25; 0.75) \\ V_7(1; 0.25; 0.75) \\ V_8(-1; 0.25; 0.75) \end{cases} \quad M = \begin{cases} V_1 \rightarrow V_2 \rightarrow V_6 \rightarrow V_5 \times C_G \\ V_2 \rightarrow V_3 \rightarrow V_7 \xrightarrow{11} V_6 \\ V_3 \rightarrow V_4 \rightarrow V_8 \xrightarrow{12} V_7 \\ V_4 \rightarrow V_1 \rightarrow V_5 \rightarrow V_8 \xrightarrow{13} V_7 \times C_M \\ V_1 \rightarrow V_2 \rightarrow V_3 \xrightarrow{14} V_4 \\ V_5 \rightarrow V_6 \rightarrow V_7 \xrightarrow{15} V_8 \end{cases}$$

Moreover, the faces of the created body are indicated using the following labels (letters) in red (see Figure 4): U stands for "Up", D stands for "Down", F stands for "Front", B stands for "Back", L stands for "Left", R stands for "Right".

Acquired raw data are integers. The acceleration values are multiplied by a constant C_A to convert them into g units, the angular velocity values by C_G to convert them into $^\circ/s$, and the magnetic field values by C_M to transform them into Gauss. These constants are equal to $2\rho/2\gamma$, where ρ is the sensor measurement range and γ is the sensor resolution. Figure 5 illustrates the flowchart of the

proposed algorithm, and the corresponding instructions and commands are detailed below. The algorithm begins by reading data, converting them, and storing them in arrays. Quaternions Q are updated using Madgwick's orientation filter, and Euler angles are then calculated, when an acceleration value is read (Type='A'). After ξ acquisitions, the 3D graph is updated by rotating the midpoints between the opposite vertices and the labels around the axes.

Algorithm 1 Real-time Orientation & 3D Visualization

```

1: Define:  $V_i, M, C_A, C_G, C_M, \xi, counter = 0$ 
2: Initialize:  $Q$  (set to 0)
3: while true do
4:   Read data-point  $(p_x; p_y; p_z)$  from the sensing device
5:   if type = 'A' then
6:      $A_x = p_x \times C_A; A_y = p_y \times C_A; A_z = p_z \times C_A$ 
7:      $Q \leftarrow$  Madgwick filter;  $\bar{Q} \leftarrow conjugate(Q)$ 
8:     Estimate  $\psi, \theta, \phi$  from  $\bar{Q}$  using (2)
9:   if type = 'G' then
10:     $G_x = p_x \times C_G; G_y = p_y \times C_G; G_z = p_z \times C_G$ 
11:   if type = 'M' then
12:     $M_x = p_x \times C_M; M_y = p_y \times C_M; M_z = p_z \times C_M$ 
13:     $counter \leftarrow counter + 1$ 
14:   if counter =  $\xi$  then
15:     $B = patch(V_i, M)$ ; Assign labels  $\eta$ 
16:     $d_1 = \frac{V_2+V_7}{2}; d_2 = \frac{V_4+V_7}{2}; d_3 = \frac{V_6+V_8}{2}$ 
17:     $rotate(B, d_1, \phi); rotate(\eta, d_1, \phi)$ 
18:     $rotate(B, d_2, \theta); rotate(\eta, d_2, \theta)$ 
19:     $rotate(B, d_3, \psi); rotate(\eta, d_3, \psi)$ 
20:     $counter = 0$ 

```

Figure 4 illustrates the orientation of the wearable device during a forward fall while walking caused by syncope (fainting). Two moments are identified: the first one (left graph) represents the orientation of the device while the subject was walking, whereas the second one corresponds to the hit on the floor.

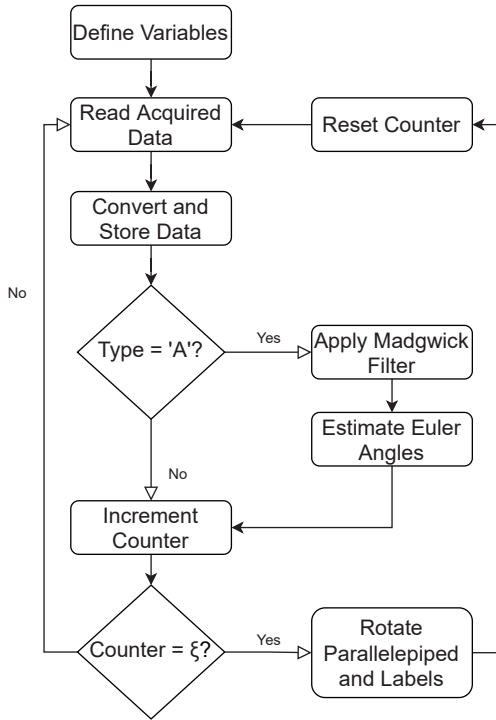


Figure 5: Flowchart for the proposed real-time orientation & 3D visualization algorithm.

4. The Proposed Strategies for Telemedicine

4.1. Towards Arm Tele-Rehabilitation

The output of the previous operations consists of three time-series, representing the sequence of Euler angles. The patterns of these signals are processed to adjust, improve, and restore the functional behavior of a subject. Tele-rehabilitation sessions are then secured based on this technique. Both the practitioner and the patient wear the same equipment (one or multiple devices, depending on the treatment) to begin the session. The former instructs the latter by performing the required movements. Hence, the clinician's signals constitute the ground truth, denoted as GT, and those of the patient are denoted as RM. Afterwards, the similarity of both signals is quantified. These measurements are important to see if the subject was able to perform the demanded task, or if he needs another session.

Two types of movement can be imposed by the practitioner, namely (i) cyclic movements, where recurrent movements are involved, and (ii) transitional movements, where a switch from an initial position to a final one occurs. In the first case, the similarity is equal to the correlation between GT and RM, as the ratio of the covariance to the product of the standard deviations. In the second case, the error is equal to the absolute value of the final position α_2 of GT minus that of RM over the change in level $\delta\alpha$ between both positions ($\alpha_2 - \alpha_1$). Hence, the similarity index v_α is:

$$v_\alpha = \begin{cases} \frac{\text{cov}(\alpha_{RM}, \alpha_{GT})}{\sigma_{\alpha_{RM}} \sigma_{\alpha_{GT}}}, & \text{for cyclic movements} \\ 1 - \left| \frac{GT \alpha_2 - RM \alpha_2}{\delta\alpha} \right|, & \text{for transitional movements} \end{cases} \quad (5)$$

It is worth mentioning that the action occurs over one or several axes. Figure 6 shows an example for wrist-worn devices. We can see that θ is exclusively involved during the transition from position (1) to position (2) in Figure 6.a, while ϕ is the only one involved in Figure 6.b. Therefore, v between θ_{GT} and θ_{RM} is computed for the first scenario (raising the arm), while v between ϕ_{GT} and ϕ_{RM} is calculated for the second scenario (rotating the arm). Note that, for other movements, two or three angles could be involved. Here, v is calculated for each affected angles (*i.e.* for each direction). Finally, a confidence interval (CI) is set by the practitioner, depending on the difficulty of the sessions/exercises and the resemblance of both signals. The assessment is based on the values of v , and whether or not they belong to CI.

4.2. Towards Human Activity Recognition

The attitude of the device (estimated Euler angles) over a period of time provides important clues in recognizing physical activities. The successive movements of the trunk should be exploited to address the problem of HAR. The considered classes are basic ADLs which are encountered during daily routine of elderly, namely (C_1) staying still, (C_2) sitting-down/standing-up, (C_3) lying-down/rising-up, (C_4) walking, (C_5) climbing/descending stairs, and (C_6) falling. A 6-second window is chosen for this task, *i.e.* a prediction is done after inspecting the movements over 6 s. The chosen length is adequate since smaller windows may exclude informa-

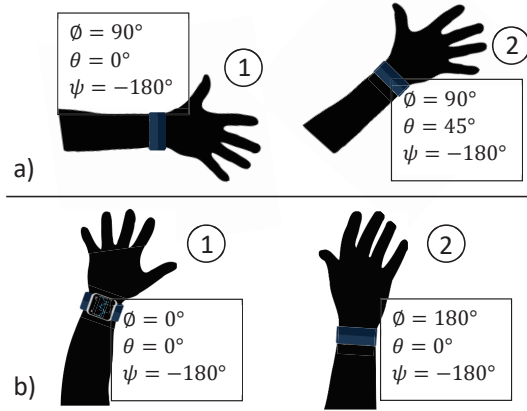


Figure 6: The Euler angles estimated from a wrist-worn device (a) while raising the arm (front view) and (b) while rotating the arm (top view).

tive parts for the recognition, while larger ones may contain several activities.

On the one hand, the recognition could be done by extracting handcrafted time-domain and frequency-domain features from raw signals, to feed classifiers like k -Nearest Neighbors (KNN), Random Forest (RF), and Deep Neural Network (DNN) [22, 30, 36]. Moreover, Random Subspace (RS) technique has been proposed to process Quaternions and Euler angles [37]. On the other hand, deep learning techniques like Convolutional Neural Networks (CNN) [9, 12], and recurrent networks like Long Short-Term Memory (LSTM) [24, 21] have been also proposed. Here, the feature extraction is done automatically.

In this paper, we present a novel recognition process. It is based on feature fusion (both handcrafted and automatically learned features) and a combination of several classification processes. The patterns of Euler angles, resulting from the sensor fusion technique (section III), feed LSTM. This network is able to learn and remember over long sequences. It supports three parallel time-series of input data, *i.e.* ψ , θ , and ϕ . This model learns an internal representation of the aforementioned sequences by extracting features automatically and mapping them to the six activity classes. It is defined by a single layer of 100 units. Now, the orientation of the trunk may not be sufficient for an accurate identification. For example, the orientation

of the human body while falling can be similar to the orientation while lying down in some situations. The speed of the movement and the shock resulting from the activity are an added value to increase the reliability of the recognition process. Such parameters are derived from the acceleration of the trunk. Hence, another deep learning module is added. It consists of a CNN and takes acceleration signals as input. CNN models usually deal with image classification problems, since they exploit the spatial correlation in data. Thus, they are appropriate to capture the shape of acceleration time-series. These signals are re-arranged to constitute an activity image of size 360×3 . Each column accounts for one component of length $6 \text{ s} \times 60 \text{ Hz}$. This model is defined by five convolutional layers, consisting of 16, 32, 64, 128, and 256 filters respectively. The filter size of the first layer is 9×3 , while the filter size of the remaining layers is 9×1 . Those layers are followed by Max Pooling layer, to reduce overfitting to the training data.

The benefit of deep learning techniques is their ability to learn automatically from raw data as mentioned before. They do not require strong expertise in feature engineering. Nonetheless, the handcrafted features complement the learned features for a better pattern knowledge. Now, these handcrafted features are extracted from both data, *i.e.* Euler angles and acceleration. Accordingly, the acceleration magnitude $\|a\|$ is first computed as $\sqrt{a_x^2 + a_y^2 + a_z^2}$. Euler angles and $\|a\|$ are split into three equal segments (2 s) without overlapping:

$$\psi_i, \theta_i, \phi_i, \|a\|_i, \text{ for } i = 1 \rightarrow 3$$

The mean values of $\{\psi_i, \theta_i, \phi_i\}$ and the standard deviation of $\|a\|_i$ are computed, resulting in 12 features. Additionally, the correlation between each pair of Euler angles $\{\psi, \theta, \phi\}$ is computed. The same goes for each pair of $\{a_x, a_y, a_z\}$. The last two operations result in 6 features. Finally, those 18 features are scaled using the sigmoid model. The scaled feature vector feeds a third unit consisting of three fully connected layers (Dense). The numbers of neurons are 40, 20, and 10 respectively. The activation function of these layers is $\hat{\sigma}$ ReLU $\hat{\sigma}$. The outputs of these three networks are concatenated, in order to feed the output layer including a softmax function which generates a probability distribution over 6 classes to predict the output of the

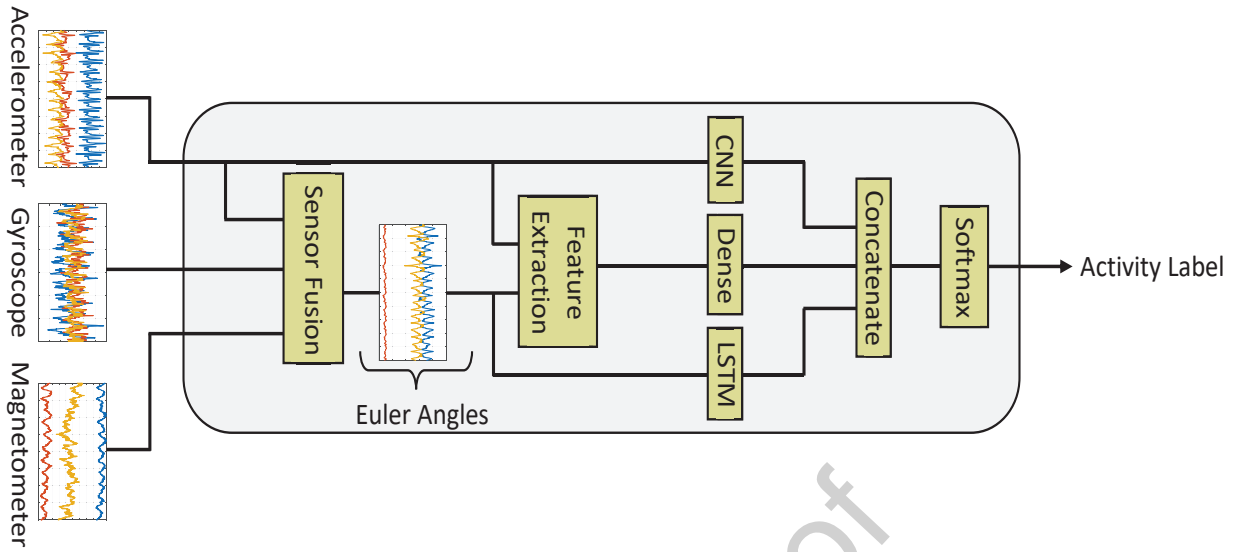


Figure 7: The proposed HAR mechanism, consisting of two deep learning networks (CNN & LSTM) and a fully connected network (dense layers).

performed ADL. Figure 7 illustrates the diagram of this identification technique.

5. Experiments

The proposed approaches were tested to validate their efficiency and prove their biomedical interest. A wrist-worn device developed by RF-Track society (located in Cesson-Sévigné, France) was used to perform some rehabilitation exercises under supervised conditions. The corresponding configuration of the device is $A F_s = G F_s = 60$ Hz and $M F_s = 40$ Hz. Moreover, we resorted to a public dataset called FallAID [29]. Three devices were considered while simulating ADLs, namely wrist-worn, waist-worn, and necklace devices. Here, $A F_s = G F_s = 238$ Hz and $M F_s = 80$ Hz. The acceleration and angular velocity signals were down-sampled by a factor of 4 (~ 60 Hz) and magnetic fields by a factor of 2 (~ 40 Hz).

5.1. Tele-Rehabilitation Module

One subject has worn the device around the wrist to test the first module. He played the role of the healthcare provider as well as the patient for an optimal evaluation. Two types of movements

Table 1

The median (MD), minimum (Min) and maximum (Max) values of the reported similarity index ν following two types of movement

Type of movement	Median	Min	Max
Cyclic	0.71	0.26	0.87
Transitional	0.76	0.49	0.99

were considered: transitional and cyclic (recurrent). He first generated a set of eight time-series (four cyclic movements and four transitional ones) as being the practitioner, by performing several rehabilitation exercises under supervised conditions. These signals were labeled as GT. In a first trial, he tried to reproduce the exact same movements (labeled as RM_1). In a second trial, he poorly replicated these gestures (RM_2). When the subject is reproducing the movements correctly (Trial 1), their patterns are close to GT. This is not the case for Trial 2. For each of the eight exercises, the similarity index was computed between GT and RM_i ($i=1 \rightarrow 2$), depending on the type of the movement.

Table 1 illustrates the median, minimum, and maximum values of the resultant eight similarity indices for each type of movement. ν ranges between

26% and 87% for cyclic movements, while that of transitional movement goes from 49% to 99%. This was expected since the second type of movement is easier to mimic. This remark is important and gives an idea regarding the choice of CI. In general, the threshold for acceptance should be lower for cyclic movements, where a similarity index around 80% could be sufficient. A video is created to show some examples regarding this module (see sub-section 5.4).

5.2. HAR Module

FallAID dataset was used in this experiment, taking both the necklace and waist-worn devices into account (body trunk). For transient ADLs, *i.e.* C_2 , C_3 , and C_6 , five 6-second windows were considered: the one centered on the transition moment, and four others shifted by ± 500 ms and ± 1 s. Note that C_2 consisted of four transitional movements, namely (i) walk-to-sit, (ii) stand-to-sit, (iii) sit-to-stand, and (iv) sit-to-walk. Moreover, C_6 consisted of falls without recovery, from three different initial body postures, namely (i) fall while walking (standing), (ii) fall while sitting, and (iii) fall from bed (lying). Now, for C_1 , the same body postures were considered, *i.e.* (i) inactive while standing, (ii) inactive while sitting, and (iii) inactive while lying. These factors are important to fairly assess the discrimination power of the proposed recognition process. As for cyclic ADLs (C_4 and C_5), five random 6-second windows were considered, since the movements were recurrent. These operations resulted in 8170 samples. This number is sufficient to train deep learning models.

10-fold cross validation was applied to evaluate the performance of our system. The networks were first trained on 90% of data (randomly chosen), using the RMSprop optimizer and the categorical cross-entropy as a loss function. The number of epochs was 12 and the batch size was equal to 64. The trained networks were then tested on the remaining 10% of data (unseen data) and the accuracy was saved. This process was repeated 10 times. The final achieved accuracy was denoted as the mean of the aforementioned ten accuracy values.

Figure 8.a illustrates the confusion matrix of the proposed model. The achieved accuracy is 97.7%. The main confusion occurs between walking and climbing/descending stairs. This was expected

since both classes are quite similar in terms of acceleration and trunk orientation. Another noticeable confusion is the one between using the stairs and sitting-down/standing-up, particularly walk-to-sit and sit-to-walk. When the subject is using the stairs, the intensity of the movements significantly decreases when he reaches the staircase landings. This pattern resembles that of C_2 to some extent. To solve this issue, the number of DOF should be increased, following two ways: either (a) by using another device located on the thigh or the ankle for example, since the movement of the leg while using the stairs differs from walking and/or sitting/standing, or (b) by mixing C_4 and C_5 into one class while adding another sensor, namely a barometer, which can measure the difference in altitude in a post-processing module. Figure 8.b illustrates the confusion matrix after mixing C_4 and C_5 . The achieved accuracy in this 5-class model increases up to 99.65%, and the confusion between classes is negligible. In this case, when the output is C'_4 , barometric data are then checked to see whether the altitude is increasing/decreasing in the sliding window (using the stairs) or not (walking). Note that this 5-class model could be sufficient for a wide-range of medical applications.

	C_1	C_2	C_3	C_4	C_5	C_6
C_1	99.36	0	0.26	0	0	0
C_2	0.42	99.1	0	0.32	2.69	0.35
C_3	0.1	0	99.74	0	0	0
C_4	0	0.06	0	96.86	5.77	0
C_5	0	0.25	0	2.82	91.54	0
C_6	0.12	0.59	0	0	0	99.65

a)

	C'_1	C'_2	C'_3	C'_4	C'_5
C'_1	99.57	0	0	0	0
C'_2	0.32	99.58	0	0.62	0.17
C'_3	0.11	0	100	0	0
C'_4	0	0.22	0	99.3	0
C'_5	0	0.2	0	0.08	99.83

b)

C'_1 : Staying still – C'_2 : Sitting down/Standing up
 C'_3 : Lying down/Rising up – C'_4 : Walking – C'_5 : Falling

Figure 8: The confusion matrix of (a) the 6-class model and (b) the 5-class model after combining C_4 and C_5 into one class.

Table 2

The experimental results comparing the computational load of orientation filters in terms of execution time (s)

Filter	EKF	Mahony	Madgwick	AQUA	Fourati
Time	1.03	0.28	<u>0.17</u>	0.26	0.82

5.3. Comparison with State-of-the-Art Techniques

In this sub-section, we will compare our approach to state-of-the-art techniques, in order to justify our choice regarding the orientation filter, and the feature fusion architecture combining CNN, Dense layers, and LSTM (CDL).

Orientation Estimation: we implemented the aforementioned orientation filters in sub-section 2.3, namely Extended Kalman Filter (EKF), Mahony, Madgwick, AQUA, and Fourati, using Python 3.6. The scripts were executed on an Intel(R) Xeon(R) Gold 5118 2.3-GHz processor with 64 GB RAM, to estimate quaternions of a human fall over 20 s. The execution time was measured in seconds. Table 2 illustrates the results. Madgwick's filter has the fastest response, with an execution time equal to 0.17 s. This filter outperforms all the others in terms of complexity, including AQUA and Mahony. This low latency ensures small frame rate or frames per second (fraps), thus a smooth 3D visualization. As expected, Kalman filter requires high computational load, hence long execution time (1.03 s in this case). The same remark holds for Fourati's filter, which has a faster response but still considerably slow compared to Madgwick's filter. This justifies our choice of using this filter in our algorithm.

HAR technique: we tested the performance of deep learning techniques when applied individually, particularly CNN and LSTM, as well as classical machine learning classifiers, namely DNN, KNN, RF, and RS. These methods were implemented using tensorflow library on Python 3.6. The input of CNN and LSTM is raw data (time-series representing Euler angles). Besides, a large set of time-domain and frequency-domain features (proposed in [22, 30, 36]) were extracted from Euler angles to feed DNN (same architecture as in [22]), KNN ($k=5$), and RF (100 estimators). Since large feature vectors could mislead the machine, they were subject to a feature selection/dimensionality reduc-

Table 3

The experimental results comparing HAR techniques in terms of accuracy (%)

Classifier	Input	Accuracy (%)
CNN	Raw data	92.67
LSTM	Raw data	91.82
CDL	Both	<u>97.7</u>
KNN	Features	95.18
DNN	Features	91.76
RF	Features	94.12
RS	Features	90.9

tion method to retain a small yet powerful subset of features, before applying the classifiers. We tested Principal Component Analysis (PCA), Stepwise Multi-linear Regression (SMR), and Recursive Feature Elimination (RFE) which uses RF in the background to recursively eliminate weaker features, and we retained the highest achieved accuracy for each pair "classifier & feature selection method". Finally, RS was fed using 14 features, namely the values of quaternions and Euler angles (7 components) over half windows (7×2). Table 3 illustrates the corresponding accuracy for each method, showing the superiority of the CDL fusion technique reaching 97.7% accuracy. As expected, the fusion between different types of features increase the discrimination power of the classification process.

5.4. Proof of Concept

This subsection illustrates the modules and functionalities of D-SORM following different scenarios, as a proof of concept. The first scenario covers the arm TR module. The ground truth signal GT and the patient's replicated signals RM are loaded via the GUI. The user (*i.e.* practitioner in this case) can manipulate RM using buttons (arrows) to see if both time-series overlap, and to calculate the similarity index ν , depending on the type of movement (transitional or cyclic). A video named "D-SORM_TR.mp4", which visualizes this manipulation, is attached to the manuscript.

The second scenario covers the replay mode of the HAR module. A fall was simulated by a subject while wearing the MARG sensing unit on his trunk. Here, the clinician can analyze this incident by visualizing the successive movements. The wearable

device is symbolized in a 3D space (Earth frame), and the predicted activity class is displayed using the proposed architecture. By using this GUI, the user can jump forward/backward and handle the speed of the visualization. A video named “D-SORM_HAR.mp4”, which visualizes this manipulation, is attached to the manuscript.

Finally, the third scenario occurs in real conditions (live mode). A subject performs some ADLs. He is wearing the device, which is connected to the PC via Bluetooth dongle. The clinician is able to monitor the subject’s gestures in real-time, with the HAR system predicting the performed activity instantaneously. A third video named “D-SORM_RC.mp4” is attached to the manuscript. It illustrates (1) the movements of the subject, (2) the real-time data acquisition, and (3) the screen of the PC, showing the GUI.

6. Conclusion

This paper proposed a digital solution called D-SORM for e-health remote monitoring. It is based on a sensor fusion technique to analyze acquired signals from wearable sensors. The orientation of the device is estimated and visualized in real-time using a GUI. This estimation, coupled with other parameters, constitutes the basis of two biomedical applications. For arm TR module, statistical features are computed from Euler angles to quantify the success rate of the patient in performing some required movements. For HAR module, a novel machine learning architecture was developed to recognize and track human motion. Euler angles and acceleration signals feed a combination of deep learning and shallow learning networks to predict the activity class.

The acceptability of D-SORM is to be considered in a future work after finalizing and releasing the first version of the application. The frequency of cyclic movements is going to be inspected in order to add it as another feature to the TR module. Since v is somewhat limited for this type of movements, their frequency will add another dimension to the clinical assessment.

Acknowledgments

This work has been supported by the French National Research Agency (ANR) in the context of the

ACCORDS project (grant n° ANR-17-CE19-0024-01). The authors would like to thank Majd Saleh for his fruitful discussions on the orientation estimation topic.

References

- [1] Baraka, A., Shaban, H., Abou El-Nasr, M., & Attallah, O. (2019). Wearable accelerometer and semg-based upper limb bsn for tele-rehabilitation. *Applied Sciences*, 9.
- [2] Bedogni, L., Di-Felice, M., & Bononi, L. (2012). By train or by car? detecting the user’s motion type through smartphone sensors data. In *Proc. of the 24th IFIP International Conference on Testing Software and Systems (ICTSS’12)* (pp. 1–6).
- [3] Bruckner, H., Kruger, B., & Blume, H. (2014). Reliable orientation estimation for mobile motion capturing in medical rehabilitation sessions based on inertial measurement units. *Microelectronics Journal*, 45, 1603–1611.
- [4] Buke, A., Gaoli, F., Yongcai, W., Lei, S., & Zhiqi, Y. (2015). Healthcare algorithms by wearable inertial sensors: a survey. *China Communications*, 12, 1–12.
- [5] Dagdeviren, C., Su, Y., Joe, P., Yona, R., Y., L., Kim, Y., Huang, Y., Damadoran, A., Xia, J., Martin, L., Huang, Y., & Rogers, J. (2014). Conformable amplified lead zirconate titanate sensors with enhanced piezoelectric response for cutaneous pressure monitoring. *Nature Communications*, 5.
- [6] Fourati, H., Manamanni, N., Afilal, L., & Handrich, Y. (2011). A nonlinear filtering approach for the attitude and dynamic body acceleration estimation based on inertial and magnetic sensors: Bio-logging application. *IEEE Sensors Journal*, 11, 233–244.
- [7] Foxlin, E. (1996). Inertial head-tracker sensor fusion by a complementary separate-bias kalman filter. In *Proc. Virtual Reality Annual International Symposium* (pp. 185–194).
- [8] Guo, L., Berglin, L., Wiklund, U., & Mattila, H. (2013). Design of a garment-based sensing system for breathing monitoring. *Textile Research Journal*, 83, 499–509.
- [9] Ha, S., & Choi, S. (2016). Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 381–388).
- [10] He, J., Bai, S., & Wang, X. (2017). An unobtrusive fall detection and alerting system based on kalman filter and bayes network classifier. *Sensors*, 17.
- [11] Jamwal, P., Hussain, S., Mir-Nasiri, N., Ghayesh, M., & Xie, S. (2018). Tele-rehabilitation using in-house wearable ankle rehabilitation robot. *Assistive Technology*, 30, 24–33.
- [12] Jiang, W., & Yin, Z. (2015). Human activity recognition using wearable sensors by deep convolutional neural networks. In *Proceedings of the 23rd ACM international conference on Multimedia*. New York, NY, USA: Association for Computing Machinery.
- [13] Kaemarungsi, K., & Krishnamurthy, P. (2004). Modeling of indoor positioning systems based on location fingerprinting. In *Proc. of the Twenty-third Annual Joint Confer-*

- ence of the *IEEE Computer and Communications Societies (INFOCOM)* (pp. 1012–1022).
- [14] Kekade, S., Hsieh, C., Islam, M., Atique, S., Khalfan, A., Li, Y., & Abdul, S. (2018). The usefulness and actual use of wearable devices among the elderly population. *Computer Methods and Programs in Biomedicine*, *153*, 137–159. doi:10.1016/j.cmpb.2017.10.008.
- [15] Kuipers, J. (1999). *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press.
- [16] Madgwick, S. (2010). An efficient orientation filter for inertial and inertial / magnetic sensor arrays. https://courses.cs.washington.edu/courses/cse466/14au/labs/14/madgwick_internal_report.pdf.
- [17] Madgwick, S., Harrison, A., & Vaidyanathan, R. (2011). Estimation of imu and marg orientation using a gradient descent algorithm. In *IEEE International Conference on Rehabilitation Robotics* (pp. 1–7).
- [18] Mahony, R., Hamel, T., & Pflimlin, J.-M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, *53*, 1203–1218.
- [19] Marins, J., Yun, X., Bachmann, E., McGhee, R., & Zyda, M. (2001). An extended kalman filter for quaternion-based orientation estimation using marg sensors. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2003–2011).
- [20] McRoberts MoveMonitor (2021). Ambulatory monitoring of physical activity for up to 14 days. <https://www.mcroberts.nl/module/physical-activity/>.
- [21] Mekruksavanich, S., & Jitpattanakul, A. (2021). Lstm networks using smartphone data for sensor-based human activity recognition in smart homes. *Sensors*, *21*.
- [22] Mishra, P., Dey, S., Ghosh, S., Seal, D., & Goswami, S. (2019). Human activity recognition using deep neural network. In *2019 International Conference on Data Science and Engineering (ICDSE)* (pp. 77–83).
- [23] Morillo, L., Gonzalez-Abril, L., Ortega Ramirez, J., & de la Conception, M. (2015). Low energy physical activity recognition system on smartphones. *Sensors (Basel)*, *15*, 5163–5196.
- [24] Murad, A., & Pyun, J. (2017). Deep recurrent neural networks for human activity recognition. *Sensors*, *17*.
- [25] Pierleoni, P., Belli, A., Maurizi, L., Palma, L., Pernini, L., Paniccia, M., & Valenti, S. (2016). A wearable fall detector for elderly people based on ahrs and barometric sensor. *IEEE Sensors Journal*, *16*, 6733–6744.
- [26] Richmond, V., Davey, S., Griggs, K., & Havenith, G. (2015). Prediction of core body temperature from multiple variables. *The Annals of Occupational Hygiene*, *59*, 1168–1178.
- [27] Roine, R., Ohinmaa, A., & Hailey, D. (2001). Assessing telemedicine: a systematic review of the literature. *CMAJ*, *165*, 765–771.
- [28] Sabatini, A. (2006). Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing. *IEEE Transactions on Biomedical Engineering*, *53*, 1346–1356.
- [29] Saleh, M., Abbas, M., & Le Bouquin Jeannès, R. (2021). Fallalld: An open dataset of human falls and activities of daily living for classical and deep learning applications. *IEEE Sensors Journal*, *21*, 1849–1858.
- [30] Shoaib, M., Bosch, S., Incel, O., Scholten, H., & Havinga, P. (2016). Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, *16*.
- [31] Smith, A., Balakrishnan, H., Goraczko, M., & Priyantha, N. (2004). Tracking moving devices with the cricket location system. In *Proc. of the 2nd International Conference on Mobile systems (MobiSYS'04)* (pp. 190–202).
- [32] Valenti, R., Dryanovski, I., & Xiao, J. (2016). A linear kalman filter for marg orientation estimation using the algebraic quaternion algorithm. *IEEE Transactions on Instrumentation and Measurement*, *65*, 467–481.
- [33] WHO (2021). Global observatory for ehealth. <https://www.who.int/observatories/global-observatory-for-ehealth>.
- [34] Xu, S., Zhang, Y., Jia, L., Mathewson, K., K.I., J., Kim, J., Fu, H., Huang, X., Chava, P., Wang, R., Bhole, S., Wang, L., Na, Y., Guan, Y., Flavin, M., Han, Z., Huang, Y., & Rogers, J. (2014). Soft microfluidic assemblies of sensors, circuits, and radios for the skin. *Science*, *344*, 70–74.
- [35] Yu, L., Xiong, D., Guo, L., & Wang, J. (2016). A remote quantitative fugal-meyer assessment framework for stroke patients based on wearable sensor networks. *Computer Methods and Programs in Biomedicine*, *128*, 100–110. doi:10.1016/j.cmpb.2016.02.012.
- [36] Zhou, B., Wang, H., Hu, F., Feng, N., Xi, H., Zhang, Z., & Tang, H. (2020). Accurate recognition of lower limb ambulation mode based on surface electromyography and motion data using machine learning. *Computer Methods and Programs in Biomedicine*, *193*.
- [37] Zmitri, M., Fourati, H., & Vuillerme, N. (2019). Human activities and postures recognition: From inertial measurements to quaternion-based approaches. *Sensors*, *19*.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Journal Pre-proof