MASTER THESIS

MINIMAL SENSING OF QUALITY OF UPPER-LIMB MOVEMENTS

Mahdad Jafarzadeh Esfahani

BIOMEDICAL SIGNALS AND SYSTEMS
PROF. DR. IR. PETER H. VELTINK

EXAMINATION COMMITTEE
Prof. dr. ir. Peter H. Veltink (Chair)
Dr. ir. Bert-Jan F. van Beijnum
Mohamed Irfan Mohamed Refai, MSc
Dr. E.H.F. Edwin van Asseldonk

DOCUMENT NUMBER
EWI-BSS – 19-24

13/08/2019

UNIVERSITY OF TWENTE.

ACKNOWLEDGMENT

I am really thankful to all the people with whom I have had the pleasure of working during my master thesis. This work and its final achievement would have not been attainable without their assistance.

I want to first extend my profound gratitude to Mohamed Irfan Mohamed Refai and Dr Bert-Jan van Beijnum, my experienced supervisors, on account of their support, constructive advice, valuable ideas, and their constant patience. They have always directed me to the right path and focus every time I found myself doubtful about parts of my research. Furthermore, I thank them for all the brilliant comments I received during our helpful progress meetings.

I would like to thank the experts who were involved in my master's thesis committee, namely: Prof. Peter Veltink and also Dr Edwin van Asseldonk due to their extensive feedback, insightful comments, and guidance that resulted in modifying some very significant sections of my report. I am grateful that they took out the time to be part of my committee.

It was an honour to have had the opportunity to work with all of the colleagues I shared my office with, especially Bouke and Miguel. Our discussions on our work gave me more clarity and encouraged me to keep persevering.

Special thanks should also go to the technicians of the lab, Ed and Marcel and also our secretary, Sandra, for helping me throughout my thesis procedure. They have always provided me with the necessary equipment and technical support.

Last but not least, I find myself deeply indebted to my parents for all that they have done for me. I cannot thank them enough for their continuous motivation, moral support and unwavering encouragement throughout my years of study. I wholeheartedly appreciate everything they have done for me.

Mahdad Jafarzadeh Esfahani

"Genius is one percent inspiration and ninety-nine percent perspiration."

-Thomas A. Edison-

To my parents; Minoo and Mehdi

Article

Minimal sensing of quality of upper limb movements

Mahdad Jafarzadeh Esfahani [1,*], Mohamed Irfan Mohamed Refai [1], Bert-Jan F. van Beijnum [1], and Peter H. Veltink [1]

[1] Biomedical Signals and Systems, University of Twente, Enschede, The Netherlands
* Correspondence: m.jafarzadehesfahani@student.utwente.nl ; Tel.: +31 64 157 81 06

Abstract: Physiotherapists supervise task-oriented recovery process of stroke patients during clinical assessment sessions; however, these patients should be also followed up after discharge from the clinic and during activities of daily living (ADL). Stroke survivors are faced with hardships while performing ADL due to their restricted arm functionality. Therefore, monitoring and subsequent assessment of the task performance during ADL would lead to efficient rehabilitation procedures. The aim of this study is to develop a stand-alone, minimal-sensing wearable system based on inertial measurement units (IMUs) capable of recognizing various ADL and assessing the quality of upper-limb movements. The proposed sensing system comprised three IMUs, one located at the sternum and two other sensors on the distal parts of the left and right forearms close to wrists. Human activity recognizer (HAR) of the study was developed using artificial neural network (ANN) in two different ways, namely: training based on one subject and testing on others as the preliminary approach, and 10-fold cross-validation as the main approach. Ten-fold cross-validation approach showed a higher accuracy of 94.02 % and 98.08% for detecting 2D-planar and 3D reaching events, respectively. The quality of the detected reaching events was assessed using various validated metrics such as smoothness, workspace area, and relative displacement of hand with respect to sternum. Feasibility of smoothness analysis using the proposed system was validated by the Xsens full-body capture system, although displacement estimations were problematic. Displacement estimates using minimal sensor setup were seen to be inconsistent between different reaching-included tasks by having errors ranging from 0.1 to 7.98 *cm* when compared to the full-body suit outcomes. Furthermore, in this study, it is proposed that smoothness analysis solely may not be a comprehensive metric for quality of movement assessment. It was seen that there exist some quite smooth movements in which relative displacement of hand and sternum are comparatively short, which is a sign for the existence of compensatory strategies. Therefore, the relative displacement of hand and sternum together with smoothness of movement complement the quality upper-limb movement assessment.

Keywords: stroke; human activity recognition; quality of movement; smoothness; upper-limb; IMU

---

1. Introduction

Stroke is the second most common cause of disability worldwide [1] and almost half of the stroke survivors suffer from long-term motor dysfunctionalities of upper-limb [2]. Objective and qualitative assessment of arm movements during functional tasks and clinical tests would lead to applying an effective rehabilitation procedure [3,4]. Furthermore, quality of movement could be assessed to differentiate between restitution or substitution of movement patterns [5].

Neurological recovery corresponds to the restoration or complete restitution of movement paradigm to the pre-morbid pattern [6]. Conversely, compensation refers to accomplishing a functional task while making use of alternative muscles or joints dissimilar to the ones that healthy people use [7]. To ease performing the functional tasks, patients with neurological deficits such as stroke use compensatory strategies to overcome their restricted range of motion [8]. In some cases when the impairment is completely irrecoverable, compensatory strategies can improve performing functional tasks and enhance independence of patients [9]. However, when the recovery is possible, these sub-conscious [7] compensatory strategies affect the quality of life [10] and could worsen the recovery process of stroke patients [11]. Although there exist clinical tests such as Fugl-Meyer for upper extremity sub-scale (FMA-UE) [12] and action research arm test (ARAT) [13] to appraise the level of impairments in the clinic, evaluation of movement quality during activities of daily living (ADL) is still challenging. To evaluate the performance of patients after discharge from rehabilitation centers, it is essential to analyze their movements during ADL [3] while taking the comfortability of the system into consideration. This raises the prominence of developing a wearable system with a minimal number of sensors which is capable of assessing the quality of movements in home environment.

Kinematic analysis of human movements using optoelectronic systems is useful to investigate the motor recovery of stroke survivors [14]; however, these systems are restricted to be utilized in a lab environment but not home. Recently, some studies showed the feasibility of utilizing inertial measurement unit (IMU) in home setting and during ADL for kinematic analysis [3,15–17]. Van Meulen et al. [15] made use of a single IMU on the forearm to estimate orientation, velocity, and position of lower arm. This study resulted in acceptable kinematic reconstructions when compared with optoelectronic reference. Moreover, a full-body IMU suit was used in [18] to investigate recovery of upper-limb movements of stroke patients in a prolonged duration, starting from 2 weeks before discharge from clinic until 4 weeks after. Their work also showed the feasibility of using IMUs to measure kinematics during ADL.

To appraise the quality of movements, it is required to first define which activities are of more interest. Admittedly, these activities should represent the discrimination between healthy and pathological movement paradigms. Reaching events exist in a wide range of ADL and comprise extension and flexion of arms. Reaching tasks could be smoothly done by the healthy people; however, restricted range of motion in the related joints of stroke survivors hinders regular movement pattern. Moreover, the difference between the normal and pathological muscle synergies can be inspected within reaching tasks. Therefore, assessing the quality of reaching tasks would lead to evaluate the quality of upper-limb movement.

Various metrics have been proposed to evaluate the movement quality of stroke survivors in home setting. Workspace area, relative displacement of hand with respect to sternum, and range of motion in vertical plane have shown to be well correlated with FMA-UE scores [19]. Quality of task-oriented movements could be also characterized in terms of smoothness, where a smooth movement corresponds to the well-coordinated movement resulted from healthy motor behavior [20,21]. Stroke patients have turbulent, unsmooth movements and it is shown that their discontinuous movements are caused by the production of many unwanted sub-movements [22,23]. However, a smooth movement comprises a few sub-movements which are fairly distributed over time [20]. Since the smoothness of movement increases after motor recovery [24], smoothness is known as a worthy metric to follow up the neuro-rehabilitation of stroke patients [20].

There are numerous measures to quantify smoothness as a metric to appraise the quality of movement. These metrics are mainly jerk-based, velocity-based, or based on arc length of movement. Well-known jerk-based metrics are namely: normalized jerk [25,26], logarithm of dimensionless jerk [20], normalized squared jerk [21], or normalized mean jerk [24]. Velocity-based measures comprise number of peak velocities [27–29], normalized average speed [24,30], and some studies regarding compositions of peaks from speed profile [31]. The methods based on arc length measures, seek for the complexity of either velocity or acceleration profiles to evaluate smoothness of movement, among which spectral arc length (*SPARC*) [20] measures smoothness of power spectra of the speed profile.

To assess the smoothness of movement, various studies utilized a different set of smoothness metrics; however, it is still not clear which of the metrics overcomes the rest. Recently, Scheltinga et al. [32] inspected a diverse set of metrics to report the useful ones during stroke recovery. They rejected the applicability of over 17 metrics due to having dimension, mathematical irrelevance of the equation, and non-reproducibility. Eventually, *SPARC*, correlation metric (CM), peak-velocity based metrics, and dimensionless jerk metrics were reported by Scheltinga et al. [32] to meet the requirements of being valuable metrics for the smoothness analysis. Balasubramanian et al. [20] also assessed various metrics for the smoothness evaluation in terms of *Validity, Consistency, Robustness* and *Sensitivity*. They proposed that an appropriate smoothness metric should be dimensionless, monotonic in response, sensitive to the changes in physiological range, and computationally affordable [20]. Based on these criteria, they showed the applicability of *SPARC* on both healthy people and subjects with neurological disorders. Acceptable sensitivity and robustness against measurements noise were also reported as the advantages of *SPARC* over other metrics.

To propose a wearable system to be used independently in a home environment, automatic activity recognition of system plays an essential role. Once the activities of interest are well classified, the assessment procedure could work more accurately. There exist some studies that developed automatic human activity recognizer (HAR) systems for different purposes such as classifying walking, sitting, climbing stairs, etc. Most of these studies made use of machine-learning-based pattern recognition algorithms to develop a HAR, namely: decision tree [3,33,34], artificial neural networks [35–37], hidden Markov model [38,39], and support vector machine (SVM) [40]. As an example, Mannini et al. [41] utilized one IMU on the ankle and one on the wrist to develop an activity classifier. They classified the activities into four classes, namely: walking, sedentary, cycling, and the rest. To perform classification, they used acceleration data and defined a feature vector consisted of mean, energy, frequency-domain entropy, and correlation features proposed by [33] and some features based on fast Fourier transform (FFT) and discrete wavelet transform (DWT) [42]. Eventually, their method reached 84.7% of accuracy for the data processed by the wrist sensor.

Some of the HAR development studies specifically focused on ADL recognition. As a case in point, the activity identification proposed by Van Meulen et al. [3] mainly focused on detecting reaching events. The method begins with identifying the posture (sitting, standing, walking) as the first step and subsequent detection of reaching task based on the hand displacement estimation. It is believed that the decision-tree approach used in this study is a primitive form of a classifier that only makes use of relative displacement between hand and sternum as the main feature. Furthermore, one can also claim that position estimation based on IMUs is not error-free and suffers from drift issues due to double integration of acceleration signal. Thus, using a larger feature set in addition to a more robust machine-learning approach could lead to having a better classification accuracy. Liu et al. [43] classified different physical activities using three sensors, namely: an accelerometer located on the distal part of left lower arm, one accelerometer on the hip, and a ventilation sensor on the abdomen to measure respiration. Their feature-set comprised 33 features that fed into SVM classifier and resulted in activity monitoring by the accuracy of 87.3 % for classifying 13 ADLs. Their proposed method is not applicable for ADL usage on account of uncomfortably of having a sensor on the abdomen throughout the day. As another example, a smart-home based approach is suggested by [44] which requires various sensor nodes in different locations of the home. Complexity of the required equipment in this work is an essential drawback for daily usage. Zhan et al. [45] also proposed a recognition system based on the environmental background sound to classify the activities by hidden Markov model (HMM) method. Their method is believed to be inconsistent between different environments and also considerably sensitive to the noise. So, automatic recognition of reaching events has only studied by [3] up to now and to the best of our knowledge there is no study on distinguishing 3D from 2D-planar reaching events during ADL. However, Kwakkel et al [46] proposed that the metrics to evaluate the 3D movements are not necessarily the same as the ones applicable to 2D tasks. Although some of the metrics are mutual between both groups, some of them are solely useful for either 2D or 3D movements. Index of curvature (IC) and peak angular velocity of elbow extension [47] are the examples of metrics which are only useful for 2D and 3D reaching events, respectively.

This study aims to develop a wearable system based on IMUs capable of assessing the quality of upper limb movements during ADL. The first step toward proposing such a system is to develop an automatic activity recognition method using IMU data to distinguish activities of interest such as reaching tasks from the rest of ADL. Afterward, the system should be able to utilize various metrics such as smoothness, relative displacement between hand and sternum, and workspace area in order to evaluate the quality of movement. As this system is designed for long measurements within ADL, comfortability or wearability of the sensing system plays an essential role. This raises the prominence of developing a system with a minimal number of sensors comprising three IMUs, one located at the sternum and two other sensors on the distal part of left and right forearms close to wrists.

2. Materials and Methods

An overview of the processing pipeline of this work is illustrated in figure 1. The work started with pre-processing of IMU data which was mainly low-pass filtering of accelerometer, gyroscope and magnetometer signals. Afterward, low-pass-filtered acceleration and angular velocity data were used to create an activity recognizer. The role of activity recognizer in this study is to mainly distinguish reaching events from the rest of ADL. This was done to be able to use automatically-detected reaching events for further analysis of movement quality in terms of various metrics such as smoothness and displacement. Moreover, Madgwick filtering [48] was used to estimate the sensor orientation in the global earth frame by fusing acceleration, angular velocity, and magnetic field data. The quaternion output of Madgwick filter was then used to transform acceleration signals from sensor frame into global earth coordinates which would lead to estimation of displacements in the global earth frame, after double integration. All required steps of the study to achieve the final aim are introduced in the following subsections.

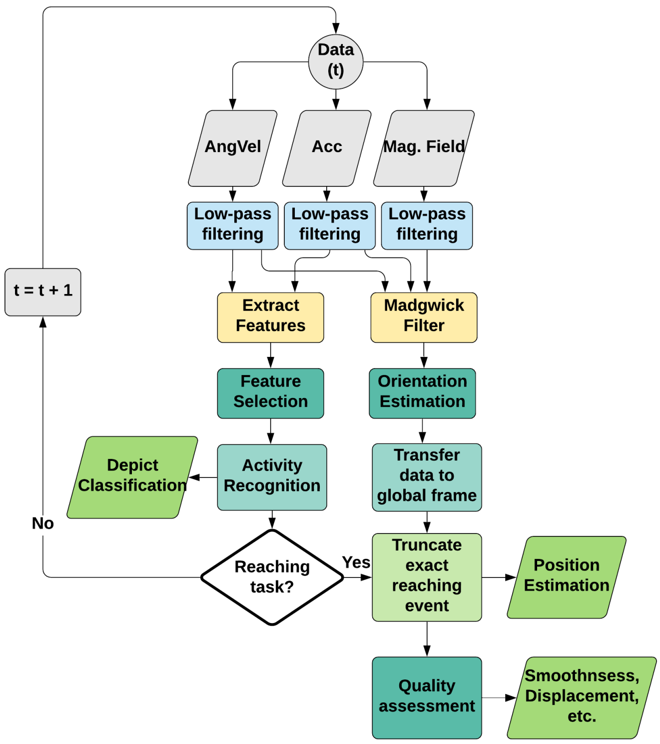

Figure 1. Overview of the processing pipeline of the study. Raw data include angular velocity (AngVel), acceleration (Acc), and magnetic field (Mag. field) received from gyroscope, accelerometer, and magnetometer, respectively.

2.1. Data pre-processing

The raw acceleration and gyroscope data were low-pass filtered by a butter-worth filter of order 8 with a cut-off frequency of 20 Hz, as suggested by [15]. All data were processed and analyzed using MATLAB® (MathWorks Inc., Natick, MA).

2.2. Activity recognition

The overall workflow of the developed HAR in this study is schematically shown in figure 2. The proposed HAR is capable of classifying input activities into 10 classes comprising 7 static poses and 3 dynamic activities, as introduced in table 1. To be able to evaluate the performance of the HAR, a reference video was recorded during experiments.

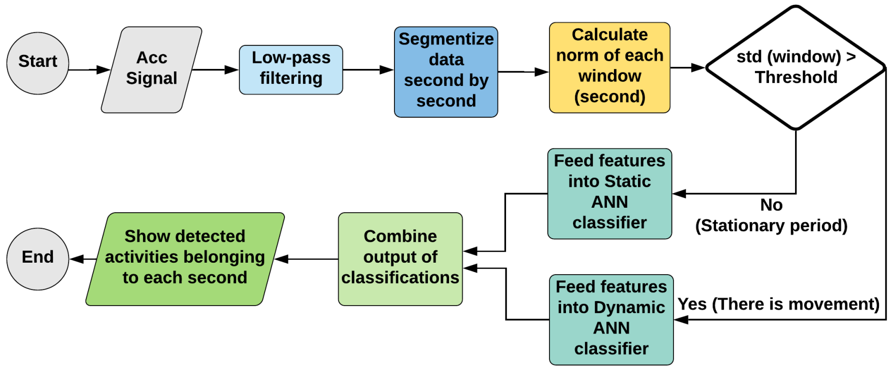

Figure 2. Workflow of the proposed HAR system. The whole dataset is firstly segmentized into windows of one second. Motion detection is based on calculating the deviation of the norm of acceleration from the nominal value of 9.81 m/s^2 which corresponds to the gravitational acceleration during quiet periods. After the distinction of stationary from dynamic time windows, the corresponding ANN models classify windows into 7 static and 3 dynamic categories as introduced in table 1.

Table 1. Definition of HAR output classes, C1 to C10. The first 7 classes belong to static poses and last three classes belong to dynamic activities. To provide a better understanding of the static poses, they are also intuitively illustrated during the experiments. Corresponding illustrations can be found in figure 8 of the section experimental protocol (2.10).

Class	Activity / Posture
1	Palm to bottom
2	Palm to top
3	Palm to left/front (Vertical)
4	Palm to right/person (Vertical)
5	Palm to left/person (Horizontal)
6	Palm to right (Horizontal)
7	Palm facing body (during normal stance posture)
8	2D-planar reaching
9	3D reaching
10	Other activities

The performance of activity recognizer method in this research is based on data in the sensor frame. This procedure started with the segmentation of data into time windows of one second. The choice of using small windows of size 1 second was due to dealing with healthy participants that could complete reaching tasks even within a second. Moreover, a smaller window size was not desirable on account of increasing computational costs. Calculating the norm of acceleration and its standard deviation during each time window determined whether a dynamic activity was being performed or the hand remained static. This procedure is called *motion detection algorithm* in this work. Threshold of motion detection was defined to be based on standard deviation and was adaptive per subject, ranging from 0.10 m/s^2 to 0.16 m/s^2 by analyzing subject-specific data. This range was found as the trade-off such that the magnitude of normal hand tremor and added noise of signal during quiet periods do not affect motion detection. *Quiet period* was defined as the period of calibration in which the subject was standing still in N-Pose. Once all the time windows were classified into the stationary and dynamic groups, corresponding features of samples throughout the time window were extracted. Subsequently, two ANN models were developed, namely: one to classify the stationary poses and the second to classify dynamic activities.

Two approaches were utilized to develop HAR systems. In the preliminary approach, data from one subject (researcher) was used to create ANN models. The proportion of train and validation sets during the "training phase" was set to be 70% and 30%, respectively. The validation proportion during the training phase was considered to investigate the generalizability of created models. Thereafter, the data of all remaining participants were used to test the performance of created models.

The second approach was to create models based on 10-fold cross-validation [49]. After the distinction of static and dynamic activities using the proposed method of motion detection, ten folds were created for each of the static and dynamic categories separately. To generate the folds, the whole set of dynamic and static features were individually shuffled into 10 folds. Subsequently, the training phase of static and dynamic models was run independently. Within each iteration of performing cross-validation, 9 folds out of 10 were utilized to train model and the remaining fold was to test the performance of trained models. The same as in the preliminary approach, the proportion of train and validation sets during the training phase of each iteration was set to be 70% and 30%, respectively. As the final step, all the classification outputs were concatenated to reconstruct the whole dataset and then compared with the target labels. The outputs of each iteration of 10-fold cross-validation can be found in details in appendix D.

The best configuration of ANN models in terms of the number of neurons in the hidden layer was found by sweeping different number of neurons and calculating classification accuracy. To avoid over-fitting, the configuration with a moderate number of neurons (10-30) and comparable accuracy with models with a higher number of neurons was chosen. The optimum structure of static ANN models in both the preliminary and 10-fold cross-validation approaches was found to have 15 neurons in the hidden layer. But, for dynamic models, the optimum configuration was seen to comprise 25 and 15 neurons in the hidden layer of the preliminary and 10-fold cross-validation approaches, respectively.

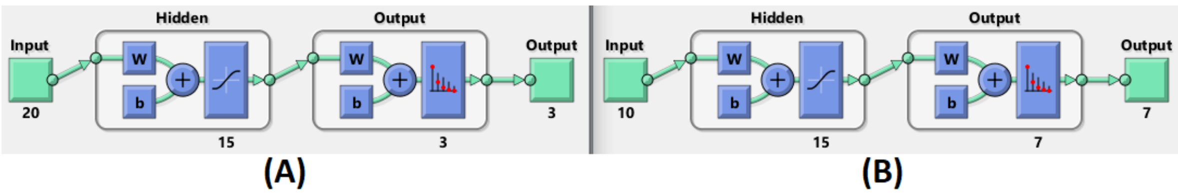

Figure 3. Architecture of the optimal configuration of ANN models using 10-fold cross-validation approach for (A) static poses and (B) dynamic activities classification. *w*: weight, *b*: bias.

To optimize model creation, different training algorithms of ANN, namely: scaled conjugate gradient (trainscg), Bayesian regularization (trainbr), and Levenberg-Marquardt (trainlm) backpropagation were tested. Consequently, for the preliminary and 10-fold cross-validation approaches, trainscg and trainlm were found as the best training techniques, respectively.

The primary list of features comprised statistical, time-domain, Fourier-domain, and wavelet-domain features adapted from some studies with a focus on HAR development [33,41,50]. There were also some features proposed by the researcher extracted from both acceleration and angular velocity signals. The proposed features by the researchers consisted of first four detail coefficients of wavelet decomposition, Shannon entropy, skewness, and kurtosis. To reduce computational costs and avoid misclassification due to having more-than-required amount of features, a feature selection method called maximum relevance minimum redundancy (MRMR) [51,52] was applied. This was to select the most essential features required for each of the static and dynamic models separately.

After MRMR-based feature selection, the static model feature set based on acceleration signal consisted of mean of axes, percentiles (25th-x, 50th-x, 25th-y, 90th-y), power in sensor z-direction, power of norm acceleration in 0.3-15 Hz frequency band, and Shannon entropy in sensor x-direction. But, the dynamic model comprised features based on both acceleration and angular velocity signals. The acceleration-based features contained mean, percentiles (25th-x, 25th-z, 25th-norm, 90th-x, 90th-z), standard deviation in sensor z-direction, first dominant frequency, powers of first three dominant frequencies, Shannon entropy in the sensor z-direction, and first wavelet detail coefficient in the x-direction. From angular velocity signal, standard deviation in sensor y-axis, 25th percentile in y-direction, and first two detail coefficient of wavelet decomposition were utilized. So, in total the static and dynamic models required 10 and 20 features, respectively.

2.3. Sensor to segment calibration

Inertial sensing systems measure data in a sensor frame of reference. However, for human motion capture, this information has to be transformed into a global frame of reference. In a sensor to segment calibration problem, the purpose is to find the relative orientation between sensor and segment $R_{Sensor}^{Segment}$. In this study, the method proposed by Choe et al. [53] was used which does not require any specific calibration movement and could be simply done in a static pose. The calibration requires standing in N-pose (or alternatively T-pose) toward north at the beginning of experiments for a few seconds. The exact direction of north was accurately defined by compass, before the start of experiments.

Due to the definition of segment coordinates of the full-body capturing system in T-pose, the calibration of the minimal sensor setup was also done in T-pose to create comparable outcomes. The first step of the calibration technique was to compensate the angle between the initial z-axis of sternum sensor and gravity such that they get aligned. After this compensation, an intermediate frame was made which was called body local coordinate system (BLCS). BLCS was created at the center of reference sensor (sternum) and X,Y, and Z axes of this frame were showing forward, left, and upward directions, respectively. The next step was to compensate the inclination error for the forearm sensors and then aligning them with BLCS. Considering the coincidence of BLCS with the global frame within calibration period, one can conclude:

$$R_{Sensor}^{Global}(qp) = R_{Sensor}^{Segment} \tag{1}$$

Where qp stands for quiet period. This way, the problem of finding segment orientation in global frame can be simplified as follows:

$$R_{Segment}^{Global} = R_{Sensor}^{Global}.[R_{Sensor}^{Segment}]^T = R_{Sensor}^{Global}.[R_{Sensor}^{Global}(qp)]^T \tag{2}$$

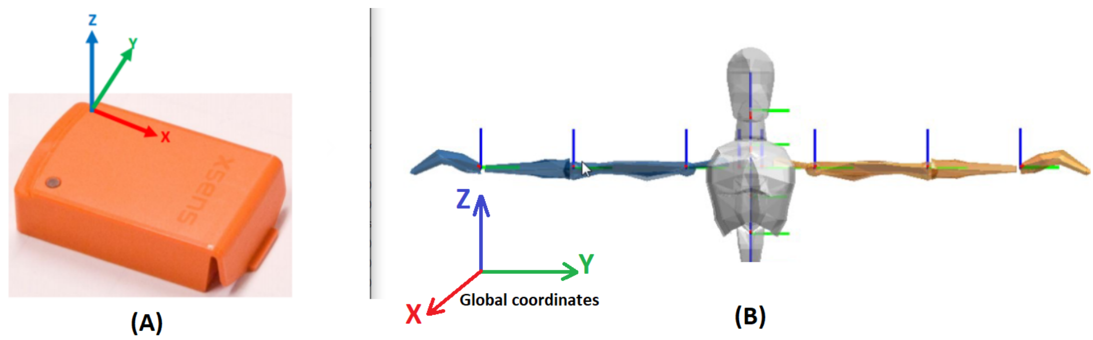

Figure 4. Depiction of (A) raw sensor frame and (B) segment frames in the end of calibration procedure in T-pose.

2.4. Orientation estimation

Finding relative orientation between the sensor and global earth coordinates in this study was based on Madgwick filtering [48]. Madgwick filter is an orientation estimation filter that benefits from low computational costs. The filter performance is based on estimation and subsequent correction of gyroscope measurement error using both accelerometer and magnetometer measurements. This approach is similar to the one proposed by Bachman et al. [54]; however, in Madgwick filter there exist much fewer computations in terms of matrix inversion and differentiation due to replacing Gauss-Newton method with gradient descent. The details of implementing filter algorithm can be found in [48]. Madgwick filter has two gain parameters that need to be defined. Filter gain β compensates the mean-zero gyroscope measurement errors that originate from different resources such as sensor noise, signal aliasing, quantization errors, and sensor axis miss-alignment [48]. Nevertheless, ζ should be set to remove gyroscope bias. In this study, these parameters were swept in the range of 0 to 0.9 by steps of 0.05 to find the best configuration. In the end, β=0.05 and ζ=0 were chosen as optimal, when compared with the reference acceleration signal in the global frame retrieved from Xsens full-body suit.

2.5. Displacement estimation

The data collected by IMU do not directly output the position information and therefore position can be estimated by the double integration of acceleration over time. The first integration of acceleration gives the velocity profile and integrating velocity would lead to position estimation. The procedure of integration is always prone to drift due to signal noise and signal offset errors [15], especially when the integration interval is longer than a few seconds [55]. In this study, the integration was only done on the corresponding time duration of reaching events which were mostly less than 2-3 seconds and therefore a simple linear compensation was applied for drift correction. Below, the step-wise procedure to estimate position from raw acceleration signal is described.

1. *Find acceleration with respect to global earth frame and remove gravity:* the raw acceleration data is in sensor frame; however, position needs to be expressed in the global earth frame. Furthermore, gravitational acceleration should be removed from z-component of acceleration to produce free acceleration.
$$\begin{bmatrix} a_x^g(t) \\ a_y^g(t) \\ a_z^g(t) \end{bmatrix} = R_s^g \times \begin{bmatrix} a_x^s(t) \\ a_y^s(t) \\ a_z^s(t) \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \tag{3}$$
Where, $a_i^s(i \in \{x, y, z\}), a_i^g(i \in \{x, y, z\})$, and R_s^g are acceleration in i-direction of sensor frame, acceleration in i-direction of global earth frame, and rotation matrix from sensor to global earth coordinates, respectively.
2. *Velocity estimation:* Using strap-down integration method, velocity in global frame was estimated as follows:
$$v^g(t) = v^g(t-1) + a^g(t) \times \Delta t \tag{4}$$
Where $v^g, a^g, \Delta t$ are velocity in global, acceleration in global, and sampling time, respectively. Since the integration was just applied on the reaching period, the initial and final conditions were defined to have zero velocity in the beginning and the end of integration period.
$$v^g(0) = v^g(end) = 0 \tag{5}$$
3. *velocity drift compensation:* Assuming a constant noise during the integration period, velocity drift could be estimated linearly.
$$\Delta v_n = \sum_{n=1}^{m} Q_n \; ; \; \forall n : \; Q_n = Q \tag{6}$$

Where Δv_n is the velocity drift at n-th sample of integration and Q stands for constant drift. For each time sample, the accumulated velocity drift over time was compensated as follows:

$$v_n = v_n - \frac{n}{m}\Delta v_m \; ; \; n = 1, \ldots, m \tag{7}$$

4. *Position estimation:* Position in global frame was estimated by integrating compensated velocity in global coordinates.

$$p^g(t) = p^g(t-1) + v^g(t) \times \Delta t \tag{8}$$

   Where p^g is the position in global frame. The initial position of hand at the predefined resting position of table was set to be $[0, 0, 0]$.
5. *Compensation of vertical displacement:* This step was only applicable on the 2D-planar reaching events that elevation was not expected to be present. To perform compensation, it was assumed that linear drift existed in the direction of vertical displacement.

$$\Delta p_n = \sum_{n=1}^{m} P_n \; ; \; \forall n : \; P_n = P \tag{9}$$

   Where Δp_n is the vertical drift of position at n-th sample of integration and P stands for constant drift. Then the corresponding portion of drift from each time sample of position estimate needed to be subtracted.

$$p_n = p_n - \frac{n}{m}\Delta p_m \; ; \; n = 1, \ldots, m \tag{10}$$

6. *Displacement estimation:* Displacement was then estimated as the Euclidean distance between the initial and final position points.

$$d = \sqrt{(x_f - x_i)^2 + (y_f - y_i)^2 + (z_f - z_i)^2} \tag{11}$$

   Where d, x, y, z were defined to be distance, x,y, and z components of position, respectively. The subscripts i and f were used to show the initial and final points, respectively. To calculate the overall displacement magnitude, the initial position point before the start of each reaching task was set to be $[0, 0, 0]$.

To perform the double integration procedure on the reaching events, the exact period of reaching should be truncated. Since the activity recognizer is segmentizing data into windows of one second, two cases might happen that influence the accuracy of displacement calculation. These situations are depicted in figure 5.A and 5.B. The first possibility is that HAR misses a small portion of data which laid on either previous or next time window (figure 5.A). The second possibility is that data in the corresponding time window comprises a long stationary period (figure 5.B). The former causes the double integration to be done on a part of reaching event but not the whole duration of reaching. However, the latter causes the integration to be applied on a larger-than-required interval which increases the drift issues.

To solve these problems, one can add a portion of both next and previous time windows to the detected reaching window to broaden the interval of detection. In this study, this portion is proposed to be the ending 70% of the prior and the beginning 70% of subsequent time windows (see figure 5.C). Based on our observations, portions of smaller size might still lose a part of reaching, while larger portions may overlap with the next/past movement event. Afterward, segmentizing the widened reaching time window into smaller windows of 0.067 seconds (15 windows per second) would lead to finding exact start and end moments of reaching (see figure 5.D). The proposed window size was found as optimum after trying different values. The same motion detection algorithm as in HAR development was used also in here to find the exact start and endpoints of reaching. In this case, threshold of motion detection was defined in the range of [0.05 - 0.1] m/s^2. This procedure was done once in the forward direction, starting from the first window to find the commencement of movement and once in the backward direction starting from the last window to find the endpoint. Once the standard deviation of the corresponding window exceeded the threshold, the window was chosen as start/end window of movement.

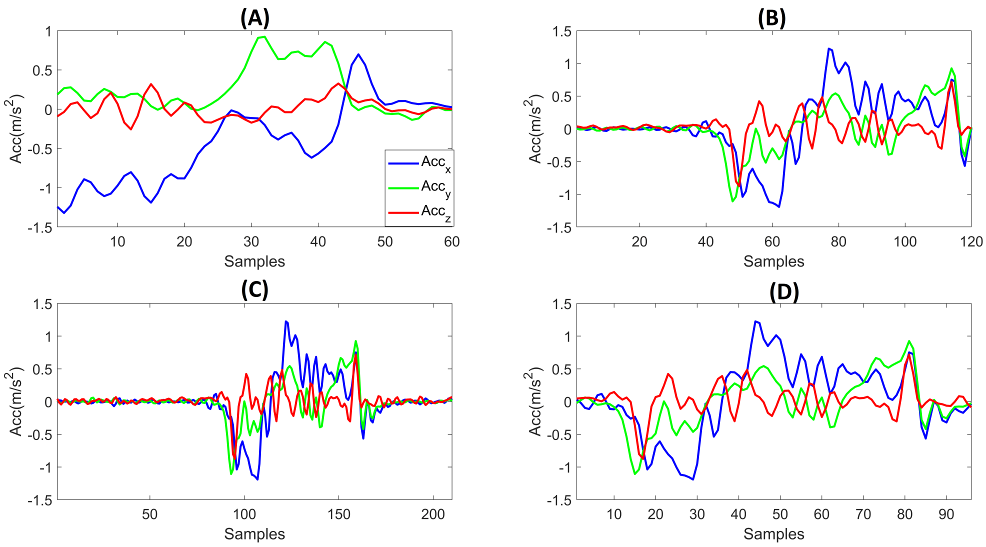

Figure 5. Sample free acceleration signals retrieved from the detected reaching events by HAR. HAR output can be in the form of (A) when a small part of signal in the beginning or end is missed, or (B) when the detected reaching event includes a large number of static samples. In (C) the same signal as shown in (B) was broadened. Then windows of size 0.067 s were applied to detect instantaneous increase/decrease in acceleration. This was done to find the exact start and endpoints of reaching event. (D) truncated version of (C) that shows accurate detection of both start and end points of movement using the proposed windowing method.

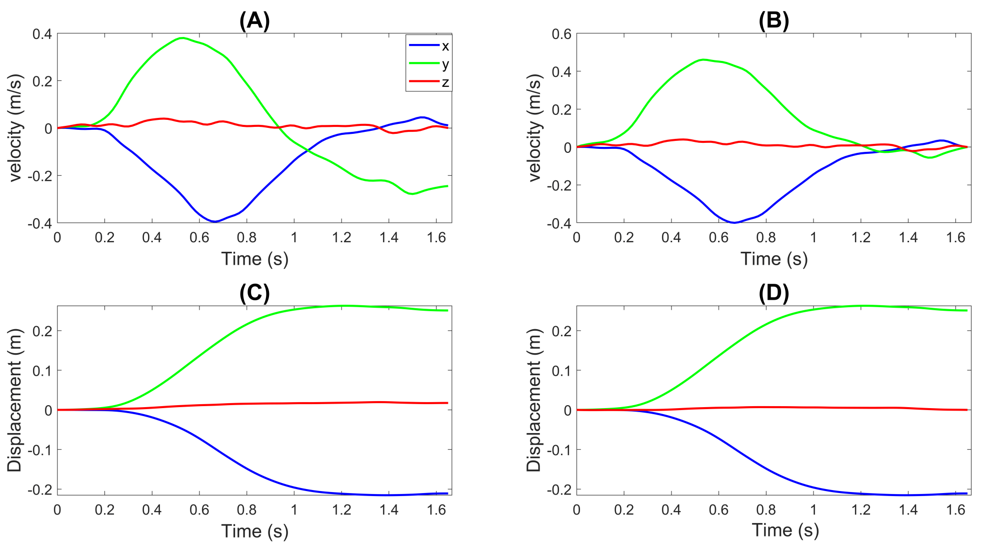

Figure 6. Displacement estimation procedure as described in section 2.5. Given an acceleration signal, in (A) the velocity profile is estimated. (B) shows velocity estimate after compensation of drift, knowing the initial and final conditions of having $v^g(0) = v^g(end) = 0$. (C) displacement estimation from the initial position point based on an estimate of the velocity profile. (D) displacement estimation after compensation of drift in elevation. Overall magnitude of displacement was then estimated as the Euclidean distance between the initial and final points in (D).

2.6. Smoothness analysis

As mentioned in the introduction section, there is no evidence of the "best" metric to quantify smoothness. However, a set of metrics which are known as the most useful ones is already reported in [32]. Furthermore, Balasubramanian et al. [20] evaluated the most relevant smoothness metrics to unify the framework of smoothness analysis.

Therefore, in this work, the developed set of metrics comprised the top ones proposed by [32] and [20]. Five different smoothness metrics were created to appraise the quality of movement, namely: $SPARC$ (both acceleration [56] and velocity [20] versions), logarithm of dimensionless jerk (LDJ) [20], number of peaks (NP) in absolute velocity profile [57], and correlation metric (CM) [58].

Velocity-based $SPARC$ [20] metric seeks for the complexity of signal in the frequency domain by the following equation:

$$SPARC_{Vel} = \int_0^{\omega_c} \sqrt{(\frac{1}{\omega_c})^2 + (\frac{d\hat{V}(\omega)}{d\omega} d\omega)} \ ; \ \hat{V}(\omega) = \frac{V(\omega)}{V_0} \tag{12}$$

Where $[0, \omega_c]$ is the frequency band, $V(\omega)$ shows the magnitude of Fourier transform of velocity profile, V_0 is the DC component of velocity profile, and $\hat{V}(\omega)$ is the normalized speed profile. In acceleration-based $SPARC$, acceleration signal replaces velocity:

$$SPARC_{Acc} = \int_0^{\omega_c} \sqrt{(\frac{1}{\omega_c})^2 + (\frac{d\hat{Acc}(\omega)}{d\omega} d\omega)} \ ; \ \hat{Acc}(\omega) = \frac{Acc(\omega)}{Acc_0} \tag{13}$$

Where $Acc(\omega)$, Acc_0 and $\hat{Acc}(\omega)$ stand for the magnitude of Fourier transform of acceleration profile, the DC component of accelration, and normalized acceleration profile, respectively.

Logarithm of dimensionless jerk metric could be calculated as follows:

$$LDJ = -ln(\frac{(t_2 - t_1)^3}{v_{peak}^2} \int_{t_1}^{t_2} \mid \frac{d^2v}{dt^2} \mid^2 dt) \ ; \ t = [t_1, t_2] \tag{14}$$

Number of peaks in velocity profile [57] is another measure to evaluate the smoothness of a signal, where a prominence of 0.05 m/s was defined in this study to find peaks.

$$NP = \{v(t) \ ; \frac{dv}{dt} = 0 \ ; \ \frac{d^2v}{dt^2} < 0\} \tag{15}$$

CM calculates the similarity of a given velocity curve with a minimum-jerk velocity profile. This ideal velocity curve [58], known as minimum-jerk, is defined as follows :

$$V_{Min.Jerk} = \Delta \, (\frac{30t^4}{T^5} - \frac{60t^3}{T^4} + \frac{30t^2}{T^3}) \tag{16}$$

Where t, Δ and T stand for time duration, trajectory distance, and the period when the velocity profile stays above a certain threshold of 1% V_{max}, respectively. The correlation between signals is a value in the interval of [0,1] that can be found as follows:

$$\rho = \frac{\Sigma[(V_{norm} - \bar{V}_{norm}).(V_{Min.Jerk} - \bar{V}_{Min.Jerk})]}{\sqrt{[\Sigma(V_{norm} - \bar{V}_{norm})^2.(V_{Min.Jerk} - \bar{V}_{Min.Jerk})^2]}} \tag{17}$$

Where ρ is correlation coefficient and $V_{norm}, \bar{V}_{Norm}, V_{Min.Jerk}, \bar{V}_{Min.Jerk}$ are normalized velocity profile, mean of normalized velocity, normalized minimum jerk velocity and mean of normalized minimum jerk velocity profiles, respectively.

2.7. Compensatory movement detection

The compensatory strategies at trunk level can be detected in this study. Detection is mainly based on measuring the displacement of the trunk during reaching events. A threshold was defined that acts as the admissible interval of displacement at trunk level. This acceptable range was found subject-specific by measuring the trunk dislocations during normal reaching tasks. Based on this method, once trunk displacement exceeds the acceptable range, the trunk is expected to perform compensation.

As shown in figure 7, the origin of sternum segment was defined in the seating position. To measure displacement, firstly the position of sternum origin (x,y,z) was converted into quaternion form:

$$u_q = 0 + xi + yj + zk \tag{18}$$

Madgwick filtering of the raw sternum sensor data led to finding the orientation of sternum in global earth frame in the form of quaternions. The resulting quaternions were then used to apply quaternion rotation operation as follows:

$$v_q = q u_q q^* \tag{19}$$

Where v_q is the rotated version of u_q vector and expresses new position of the rotated point (x,y,z) after conversion from quaternions into $\mathbb{R}^3$ space.

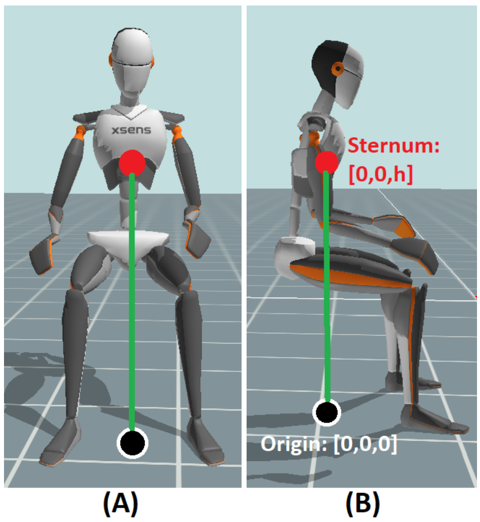

Figure 7. Sternum position with respect to the origin of coordinate system in (A) front and (B) side views. During neutral seating pose, trunk is coincident with the origin in x-y (north-west) plane but has the height h which was measured during experiments. The red circle shows the origin of sternum segment $[0, 0, h]$ and the black circle is the origin of the coordinate system $[0, 0, 0]$.

2.8. Equipment

Kinematic data in this study were measured simultaneously using two inertial sensing systems. The proposed minimal sensing system of the study comprised three Xsens Awinda IMUs (Xsens Technologies, Enschede, the Netherlands) placed on sternum and both right and left forearms, close to the wrists. Furthermore, Xsens Awinda full-body capture system comprising 17 IMUs was used as the ground truth of the research. Full-body capturing system requires sensors to be placed on head, sternum, pelvis, both sides of shoulder, upper arms, forearms, hands, upper legs, lower legs, and feet. Data were recorded simultaneously on both systems at a frequency of 60 Hz. Full-body capturing system provides accurate estimations of body segments kinematics and is already validated for kinematic analysis when compared with optoelectronics. Therefore, this system was defined as the ground truth to compare kinematic outputs of the proposed minimal sensing system with its outcomes.

2.9. Subjects

There were six healthy participants (22.0 ± 2.1 years old, 171.0 cm ± 7.8 cm height, 70.2 cm ± 2.5 cm arm length from shoulder to finger tips, 5 right-handed and 1 left-handed, 4 males and 2 females) recruited for this study. Participants had no history of neurological disorders and were provided with informed consent before the commencement of the experiment. Ethical approval was obtained from the Ethics Committee of the University of Twente.

2.10. Experimental protocol

The experiments of this study comprised two main parts, namely *static hand poses* and *dynamic hand movements*. In the beginning, subjects were asked to start imitating the items of the static poses list in order. As shown in figure 8, the list of hand poses comprised the following items: (1) palm to the bottom, (2) palm to the top, (3) palm to the left-vertical[i], (4) palm to the right-vertical[ii], (5) palm to the left-horizontal[i], (6) palm to the right-horizontal[ii], (7) palm to the front-vertical, (8) palm toward person-vertical, (9) hands parallel to body when palm facing body (normal stance posture) and (10) palm toward person-horizontal.

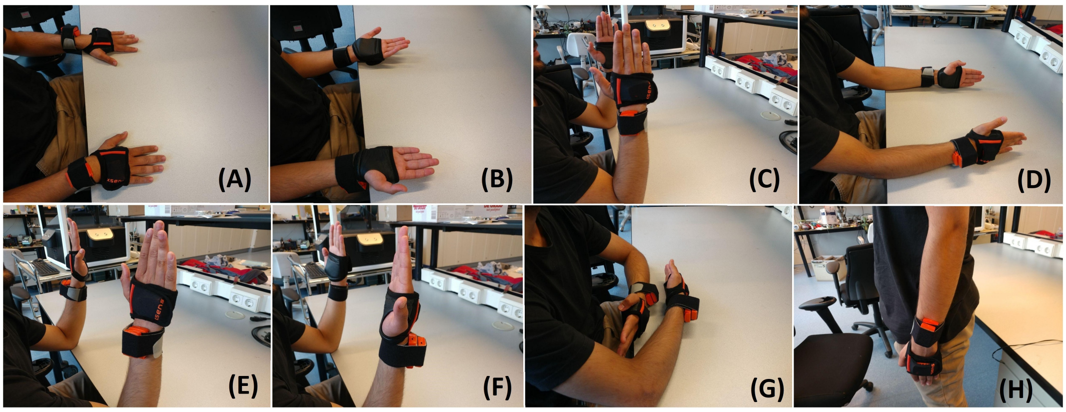

Figure 8. Various static poses of hands in the study, where each pose can be related to a specific hand pose during ADL. (A) Palms to the bottom, (B) Palms to top, (C) Palm of left hand to the right and palm of right hand to the left -*Vertical*, (D) Palm of left hand to the right and palm of right hand to the left -*Horizontal*, (E) Palms to the front, (F) Palms toward person - *Vertical*, (G) Palms toward person - *Horizontal* (H) Hands parallel to body during stance.

At the end of performing static poses, subjects were provided with the list of dynamic activities. The list of dynamic activities comprised the following:

1. Sitting reach: Participants were asked to reach their hand from the initial hand position on the table to the pre-defined location of targets, under different configurations as described below. The placement of targets is shown in figure 9.
   (a) Normal reaching: Reach targets normally without any restriction.
   (b) Extended reaching: Targets were placed 10 cm farther than their initial positions. Subjects were allowed to bend toward targets while performing the reaching task.
   (c) 2D-planar reaching on hand support (HS): Reaching was restricted to be in horizontal plane to avoid elevation. Therefore, a Hankamp MAS Mini hand support (see figure 10) was used in which elevatory degree of freedom was restricted by a metal bar (see figure 10.C). This was done to ensure the movement is in 2D transverse plane.

---

i Applicable for right hand.
ii Applicable for left hand.

(d) Extended 2d-planar reaching on HS: The same as in item 1.b, targets were placed 10 cm farther than their initial positions. But in this case, HS was used to ensure that the movement is in horizontal transverse plane.

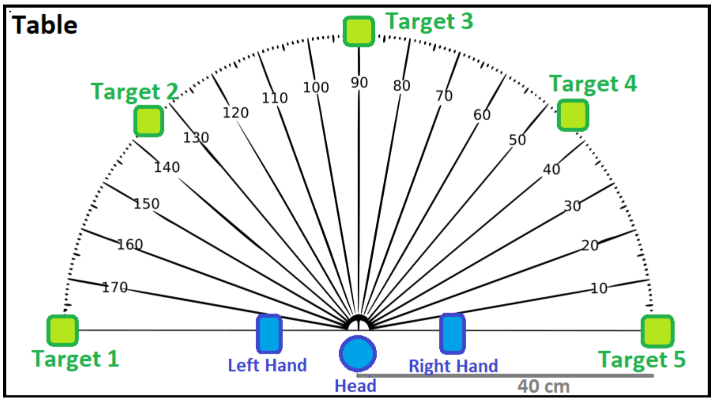

Figure 9. Placement of targets on the table for sitting reach tasks. Target 1 and 5 were aligned with hands initial position. Subsequent targets were placed at 45° deviations. Targets 1 to 4 were to be reached by left hand and targets 2 to 5 were to be reached by right hand. During extended reaching tasks, targets were placed 10 cm farther such that reaching without bending over targets was not possible.

Figure 10. Hand support in different postures. (A) when the arm of the device is closed and (B) when the arm of HS is fully extended. (C) the diagonal metal bar was added to the device to restrict the elevatory degree of freedom and limit movements to be in horizontal transverse plane.

2. 3D reaching: In this task, subject was asked to perform reaching during stance. Reaching events started from normal hand pose during stance (see figure 8.H) to reach toward targets placed on a shelf. Three targets were placed on a shelf with a height approximately at chest level of participants. The middle target was placed exactly in front of the participant during initial stance pose and two other targets were placed at 45° and -45° deviations, respectively (see targets 2 to 4 in figure 9).
3. Walking: Subject was requested to walk with the desired speed in the measurement environment, where the walking distance was approximately 13 m. The subject was asked to start from the initial position, walk toward the end of the route, turn back, and come to the initial starting position. There was no restriction on the arm movement during the task.
4. Determining the potential hand workspace: Subject was asked to perform a hand semi-circular movement in both clockwise and counter-clockwise directions. At the end of the movement, task was defined to be completed once right and left hands reached targets 1 and 5 of figure 9, respectively. Participant was asked to do this activity on both the HS and table.
5. Drinking: Subject was asked to commence with 2D-planar reaching to grasp the cup of water (plastic cup with 5-6 cm diameter, 8-9 cm height), elevating it, drinking water, put the cup on the table and come back to initial hand position on table.

6. Writing: During this task, subject was asked to rewrite a given 3-line text in the notebook. there was no restriction on the speed of writing.

The order of performing both static and dynamic tasks was randomized for each subject to guarantee the independence of HAR performance from the order of occurrence of events. Each static pose was asked to be held for 30 seconds. The frequency of performing dynamic activities was 5 times and after each moving task, there was a 5 second of static hand pose at resting position of the table (as shown in figure 8.A). This was done to ensure the possibility of distinguishing different portions of movement, e.g.: flexion from extension during reaching. The total duration of each experiment was approximately 1 hour.

2.11. Methods for analysis of results

The results section is commenced with the outcomes of the preliminary and 10-fold cross-validation approaches of HAR development. Corresponding confusion matrix and statistical analysis of each approach are reported to show the classification performance. Afterward, assessment of the quality of movement begins with an estimation of hand displacement with respect to the sternum during various reaching-included tasks. Estimated displacements by the proposed minimal sensing setup are compared with the full-body suit outcomes as the ground truth. Furthermore, the potential workspace area of hand during the correspondent workspace task is reported.

Quality of movement assessment is then followed by the smoothness analysis utilizing various metrics. Smoothness of different tasks of the study in which reaching events have existed is assessed using the retrieved data from the minimal sensor setup and full-body suit separately. Two-sample t-test is then used as an inferential statistical technique to check the potential of statistically significant difference between the means of the measured smoothness outcomes by two systems of the study, namely: minimal sensing and full-body suit systems. Eventually, estimated results of trunk displacement while performing normal and extended reaching tasks are reported. This measure is a prominent clue to inspect the absence or presence of compensatory strategies at trunk level.

3. Results

A sample output of the proposed HAR system is shown in figure 11, where the performance of HAR is visually compared with the actual reference labels of the recorded data. In this visualization method, one can determine each second of the measurement belongs to which of the defined activities.

As earlier discussed in section 2.2, two approaches were utilized to develop HAR, namely: *Preliminary* and *10-fold cross-validation* approaches. To appraise the performance of these methods, it is essential to look into different statistical measures such as accuracy, precision, sensitivity, specificity, and F1-score. Corresponding statistical analysis of the preliminary approach is represented in table 2.

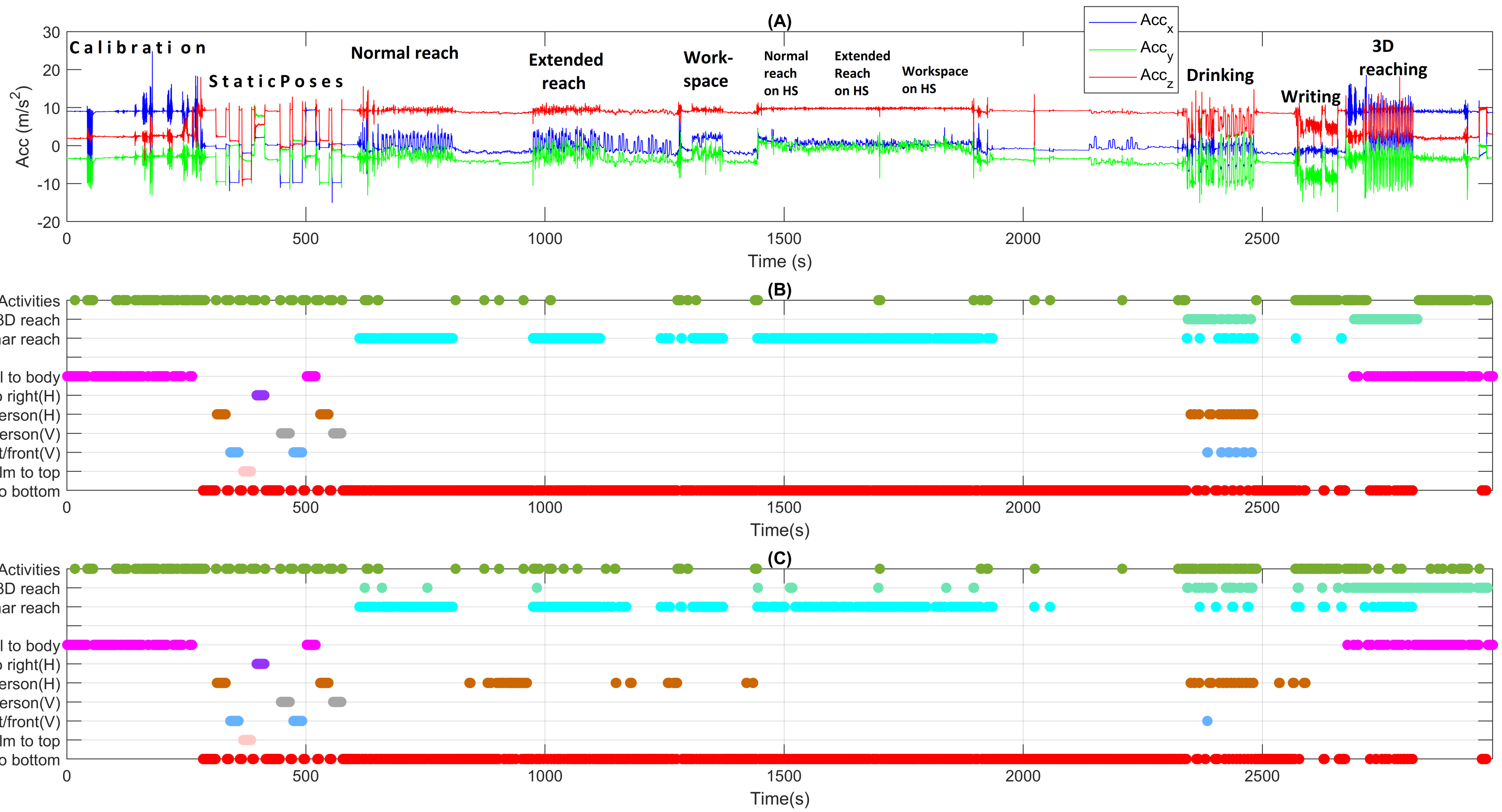

Figure 11. An example of HAR output from activities performed by a participant. (A) Raw acceleration signal retrieved from the sensor on the right forearm. Different portions of the signal were labeled to clarify the performing activity. (B) Corresponding reference labels of data (C) Classification result from HAR output.

Table 2. Statistical analysis of the classification performance of the preliminary approach of HAR development on test data. C_i ($i \in \{1, .., 10\}$) stand for classes as introduced in table 1, C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose), C8: 2D-planar reach, C9: 3D reach, C10: Other Activities. Std: standard deviation.

Class	Accuracy(%)	Precision(%)	Sensitivity(%)	Specificity(%)	F1-score
C1	93.89	95.81	94.18	93.43	94.99
C2	99.75	100.00	61.68	100.00	76.30
C3	99.72	95.36	88.70	99.92	91.91
C4	99.74	86.04	94.09	99.81	89.88
C5	97.15	39.19	94.06	97.21	55.33
C6	99.98	96.63	100.00	99.98	98.29
C7	99.04	98.91	93.22	99.86	95.98
C8	93.84	78.06	71.11	97.12	74.42
C9	96.03	43.57	54.48	97.50	48.42
C10	94.64	78.65	76.61	97.13	77.62
Mean-Dynamic classes	94.84	66.76	67.40	97.25	66.82
Std-Dynamic classes	1.11	20.08	11.52	0.22	16.02
Mean-Static classes	98.47	87.42	89.42	98.60	86.10
Std-Static classes	2.24	21.74	12.66	2.49	15.36

In table 2, an overall statistical interpretation of the preliminary approach of HAR development is given in terms of various measures; however, there is no clue about the confusion amongst classes. Thus, to interpret the misclassification between classes, corresponding confusion matrix of the preliminary approach is represented in table 3.

Table 3. Confusion matrix of the target (labels) vs. output of the preliminary approach of HAR development. Correctly classified observations are shown in diagonal cells and incorrect predictions are shown in off-diagonal cells. Both the number and percentage of the total amount of observations per class are shown in each cell and overall accuracy is represented in the bottom-right corner cell. C_i ($i \in \{1, .., 10\}$) stand for classes as introduced in table 1, C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose), C8: 2D-planar reach, C9: 3D reach, C10: Other Activities.

Output \ Target	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	
C1	**10223(54.2%)**	0	4(0.0%)	0	1(0.0%)	0	1(0.0%)	410(2.2%)	1(0.0%)	30(0.2%)	
C2	0	**66(0.3%)**	0	0	0	0	0	0	0	0	
C3	1(0.0%)	0	**267(1.4%)**	12(0.1%)	0	0	0	0	0	0	
C4	0	9(0.0%)	21(0.1%)	**191(1.0%)**	0	0	0	0	0	1(0.0%)	
C5	425(2.3%)	32(0.2%)	0	0	**301(1.6%)**	0	0	0	1(0.0%)	9(0.0%)	
C6	2(0.0%)	0	0	0	0	**86(0.5%)**	0	0	0	1(0.0%)	
C7	0	0	0	0	0	0	**1912(10.1%)**	0	3(0.0%)	18(0.1%)	
C8	167(0.9%)	0	0	0	0	0	0	**1583(8.4%)**	76(0.4%)	202(1.1%)	
C9	13(0.1%)	0	0	0	1(0.0%)	0	110(0.6%)	57(0.3%)	**322(1.7%)**	236(1.3%)	
C10	24(0.1%)	0	9(0.0%)	0	17(0.1%)	0	28(0.1%)	176(0.9%)	188(1.0%)	**1628(8.6%)**	
											87.88%

After analyzing the outcomes of the preliminary approach of HAR development, the classification output of the 10-fold cross-validation approach was evaluated. Therefore, tables 4 and 5 correspond to the statistical analysis and confusion matrix of the 10-fold cross-validation approach, respectively.

Table 4. Statistical analysis of the classification performance of the 10-fold cross-validation approach of HAR development on test data. Std: standard deviation. C_i $(i \in \{1, .., 10\})$ stand for classes as introduced in table 1. C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose), C8: 2D-planar reach, C9: 3D reach, C10: Other Activities.

Class	Accuracy(%)	Precision(%)	Sensitivity(%)	Specificity(%)	F1-score
C1	96.66	96.66	97.87	94.79	97.26
C2	99.97	94.69	100.00	99.97	97.27
C3	99.91	97.65	97.00	99.96	97.32
C4	99.98	100.00	98.04	100.00	99.01
C5	99.79	96.76	91.72	99.94	94.17
C6	99.94	91.58	97.75	99.95	94.57
C7	99.08	99.58	92.53	99.95	95.93
C8	94.02	74.16	77.69	96.27	75.89
C9	98.08	75.20	63.73	99.27	68.99
C10	95.07	78.96	79.11	97.20	79.04
Mean-Dynamic classes	95.72	76.11	73.51	97.58	74.64
Std-Dynamic classes	2.11	2.52	8.50	1.54	5.14
Mean-Static classes	99.33	96.70	96.42	99.22	96.50
Std-Static classes	1.22	2.90	3.08	1.96	1.71

Table 5. Confusion matrix of target vs. output of 10-fold cross-validation approach of HAR development. Correctly classified observations are shown in diagonal cells and incorrect predictions are indicated in off-diagonal cells. Both the number and percentage of the total amount of observations were shown in each cell. Overall accuracy is shown in the bottom-right corner cell. C_i $(i \in \{1, .., 10\})$ stands for classes as introduced in table 1. C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical) ,C4: Palm to right/person (Vertical), C5: Palm to left/person(H), C6: Palm to right(H), C7: Hands parallel to body (during stance) , C8: 2D-planar reach, C9: 3D reach, C10: Other Activities.

Target / Output	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	
C1	**10607(56.2%)**	0	0	0	9(0.0%)	2(0.0%)	0	329(1.7%)	0	27(0.1%)	
C2	3(0.0%)	**107(0.6%)**	0	0	0	0	0	3(0.0%)	0	0	
C3	1(0.0%)	0	**291(1.5%)**	4(0.0%)	0	0	0	2(0.0%)	0	0	
C4	0	0	0	**200(1.1%)**	0	0	0	0	0	0	
C5	9(0.0%)	0	0	0	**299(1.6%)**	0	0	1(0.0%)	0	0	
C6	7(0.0%)	0	0	0	0	**87(0.5%)**	0	1(0.0%)	0	0	
C7	0	0	0	0	0	0	**1895(10.0%)**	5(0.0%)	0	3(0.0%)	
C8	182(1.0%)	0	1(0.0%)	0	1(0.0%)	0	0	**1731(9.2%)**	90(0.5%)	329(1.7%)	
C9	2(0.0%)	0	4(0.0%)	0	10(0.1%)	0	9(0.0%)	12(0.1%)	**376(2.0%)**	87(0.5%)	
C10	27(0.1%)	0	4(0.0%)	0	7(0.0%)	0	144(0.8%)	144(0.8%)	124(0.7%)	**1689(9.0%)**	
											91.61%

Assessment of quality of upper-limb movement begins with displacement estimation. Relative displacement between hand and sternum indicates the reaching distance of hand while considering dislocation of trunk. Table 6 shows the comparison of displacement estimation between the proposed minimal sensor setup and full-body suit, as the ground truth. In this table, displacement estimations of less than 10 cm were excluded from the results.

Table 6. Estimation of relative displacement of hand with respect to sternum during reaching events. Displacements were calculated by both the minimal sensing and full-body suit systems to perform comparison of outcomes. Results are given in the form of *mean* $\pm$ *Std*.

Task	Ref. relative displacement (cm)	HAR relative displacement (cm)	Absolute mean error (cm)
Normal reach - HS	26.84 ± 10.24	18.87 ± 8.72	7.97
Extended reach - HS	26.77 ± 10.41	20.84 ± 7.79	5.93
Normal reach - table	28.05 ± 8.47	27.95 ± 8.02	0.10
Extended reach - table	28.15 ± 9.46	32.49 ± 11.45	4.34
Reach to grasp	27.08 ± 6.91	21.60 ± 5.06	5.48

As introduced in the experimental protocol of section 2.10, workspace task comprised hand semi-circular movements in both clockwise and counterclockwise directions to determine the potential reachable area of hands. Figure 12.A illustrates the displacement of both left and right hands during workspace task. Despite hand displacement, correspondent displacement of the trunk is also depicted while performing this task. Creating an envelope covering the lower and upper limits of the reachable area determines the potential workspace, as represented in figures 12.B and 12.C for the right and left hands, respectively. The estimated workspace area over all 6 participants of the study was found to be $0.41\,\mathrm{m}^2 \pm 0.11\,\mathrm{m}^2$ for the right hand.

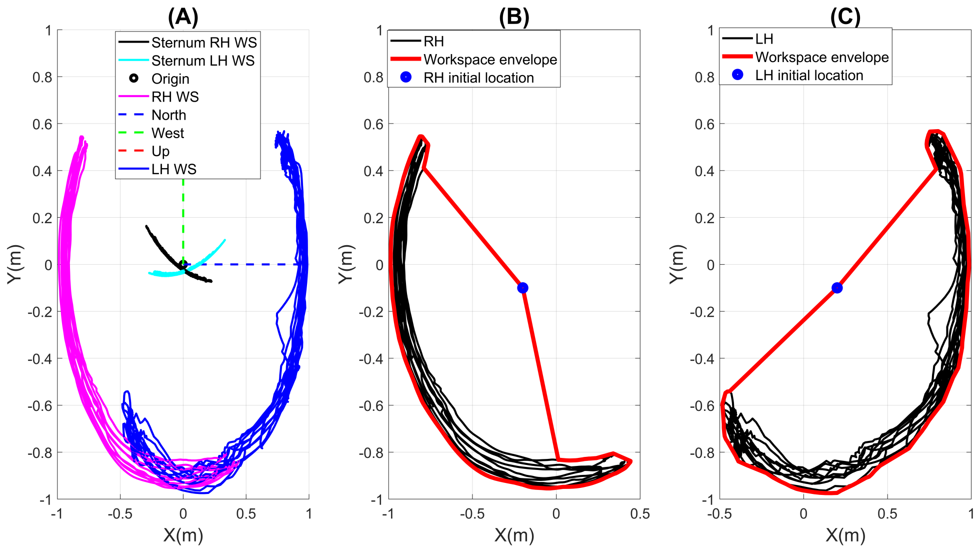

Figure 12. Determination of the potential hand workspace area. (A) Displacement of both hands and corresponding trunk dislocation during workspace determination task on table, (B) The envelope covering the area of right hand potential workspace, (C) The envelope covering the area of left hand potential workspace. RH: right hand, LH: left hand, WS: workspace.

Smoothness analysis of the reaching tasks in this study was done by various metrics as already introduced in section 2.6. 2D-planar reaching events were existed in different trials of the experiment, namely: normal 2D-planar reach-to-point on HS and table, extended 2D-planar reach-to-point on HS and table, and eventually reach-to-grasp as a part of the drinking task. Therefore, in table 7, the task-specific smoothness analysis derived from both the proposed and reference systems is shown. In table 7, CM outputs of less than 30% were extracted.

Table 7. Smoothness analysis of the 2D-planar reaching events measured by the developed HAR vs. full-body suit in terms of various metrics. Smoothness values are given in the format of $Mean \pm Std$. $SPARC$: Spectral arc length, LDJ: Logarithm of dimensionless jerk, CM: Correlation metric, NP: number of peaks in velocity profile.
*: Not applicable. Since reach-to-grasp is not a symmetric movement, it is not comparable with a minimal jerk profile. $^+$: NP is given as a proportion of the number of velocity profiles with only one peak divided by the total number of all velocity profiles. †: There is a statistically significant difference between the output of HAR and Ref (significance level of 95%).

Task	**System**	$SPARC_{Vel}$	$SPARC_{Acc}$	**LDJ**	**CM**	NP^+
Normal reach on HS	HAR	-1.65 ± 0.31	$-2.41 \pm 0.66^\dagger$	$-7.47 \pm 1.79^\dagger$	78.50 ± 18.68	75.76
	Ref.	-1.74 ± 0.66	-2.83 ± 0.84	-8.14 ± 2.40	78.27 ± 17.31	81.21
Extended reach on HS	HAR	-1.71 ± 0.33	$-2.61 \pm 0.67^\dagger$	$-8.41 \pm 1.86^\dagger$	$76.33 \pm 19.50^\dagger$	70.14
	Ref.	-1.70 ± 0.70	-2.89 ± 0.78	-8.97 ± 2.31	83.78 ± 12.47	79.58
Normal reach on table	HAR	$-1.62 \pm 0.53^\dagger$	$-2.56 \pm 1.05^\dagger$	$-7.87 \pm 1.47^\dagger$	$77.93 \pm 16.40^\dagger$	70.61
	Ref.	-1.88 ± 1.13	-2.93 ± 1.38	-8.50 ± 1.88	71.36 ± 16.28	81.91
Extended reach on table	HAR	-1.57 ± 0.36	$-2.48 \pm 0.84^\dagger$	$-8.28 \pm 1.43^\dagger$	74.28 ± 19.59	64.49
	Ref.	-1.69 ± 0.79	-2.67 ± 1.03	-8.83 ± 1.59	76.71 ± 15.30	86.80
Reach-to-grasp	HAR	-1.49 ± 0.09	-2.17 ± 0.59	-7.65 ± 1.27	N/A*	96.67
	Ref.	-1.62 ± 0.56	-2.37 ± 0.81	-8.52 ± 3.15	N/A*	85.19
Overall 2d reaching events	HAR	-1.62 ± 0.40	$-2.49 \pm 0.83^\dagger$	$-7.98 \pm 1.63^\dagger$	76.71 ± 18.50	71.72
	Ref.	-1.75 ± 0.84	-2.79 ± 1.05	-8.60 ± 2.16	77.24 ± 16.06	82.84

As the next step, two-sample t-test with a significance level of 95% was applied to explore any statistically significant difference between the smoothness outcomes resulted from the measurement systems, namely: minimal sensing setup and full-body suit. Within NP results, no statistically significant difference was found. But, two-sample t-test showed that there is a statistically significant difference (p-value= 0.00) among the outcomes of $SPARC_{Vel}$ while performing normal reach on table. The outcomes of both $SPARC_{Acc}$ and LDJ were always significantly different between systems (all by p-values=0.00) except for the reach-to-grasp. CM was seen to have significant differences during extended reaching events on HS and normal reach on table (both by p-values=0.00). Overall, there was a statistically significant difference in the mean of $SPARC_{Acc}$ and LDJ measurements, whereas no significant difference was found in the mean of $SPARC_{Vel}$, NP, and CM outputs. The detailed results of two-sample t-test can be found in appendix F.

Correlation metric is only applicable on the symmetric velocity profiles, e.g: reach-to-point. However, asymmetric velocity profiles such as reach-to-grasp events cannot be evaluated using this metric. Therefore, figure 13 depicts the boxplot of correlation metric corresponding to the symmetric reaching tasks of the study.

During extended reaching tasks that require bending over targets, one can inspect the existence of compensatory movements at trunk level. As illustrated in figure 14.A, presence of compensation led to having larger displacement and rotation of sternum during extended reaching events when compared with normal reaching. Moreover, in figures 14.B and 14.C corresponding displacement and rotation of the figure 14.A were further analyzed. Trunk displacement was seen to be 17.70 cm ± 5.00 cm and 17.62 cm ± 5.59 cm within the extended reaching tasks on HS and table, respectively. However, in normal reaching events, trunk displacement was approximately 6.54 cm ± 2.04 cm. So, the presence of compensatory strategies during extended reaching tasks is apparent.

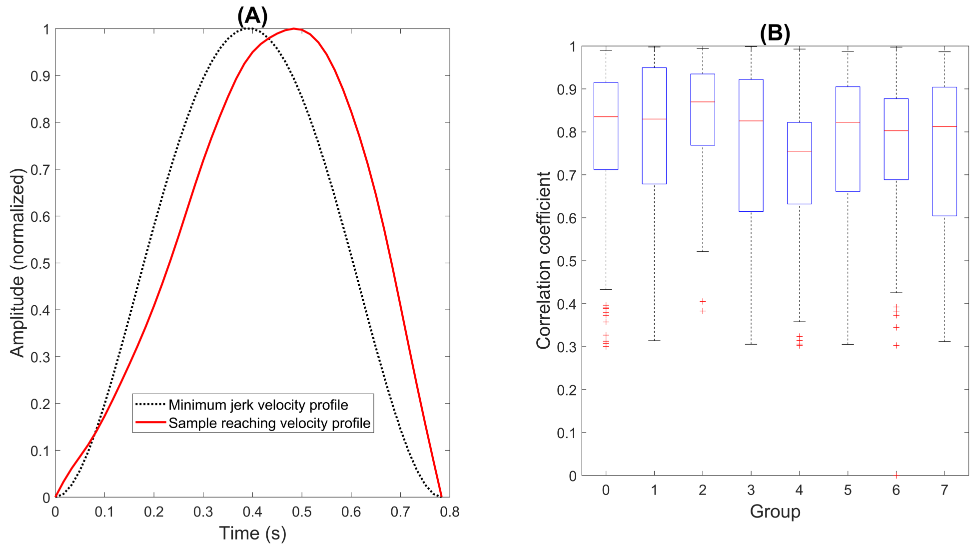

Figure 13. Smoothness evaluation of the detected reaching events using correlation metric. For each task, CM value was calculated based on the data retrieved from the minimal sensing system (HAR) and full-body suit (Ref) separately. (A) sample velocity profile of reaching task derived from the integration of acceleration profile in comparison with a minimum jerk profile produced by equation 16. (B) Box-plot of correlation metric belonging to different groups of reaching events. Groups: (0) Normal reach on HS - Ref., (1) normal reach on HS - HAR , (2) Extended reach on HS - Ref., (3) Extended reach on HS - HAR, (4) Normal reach on table - Ref., (5) normal reach on table - HAR, (6) Extended reach on table - Ref., (7) Extended reach on table - HAR.

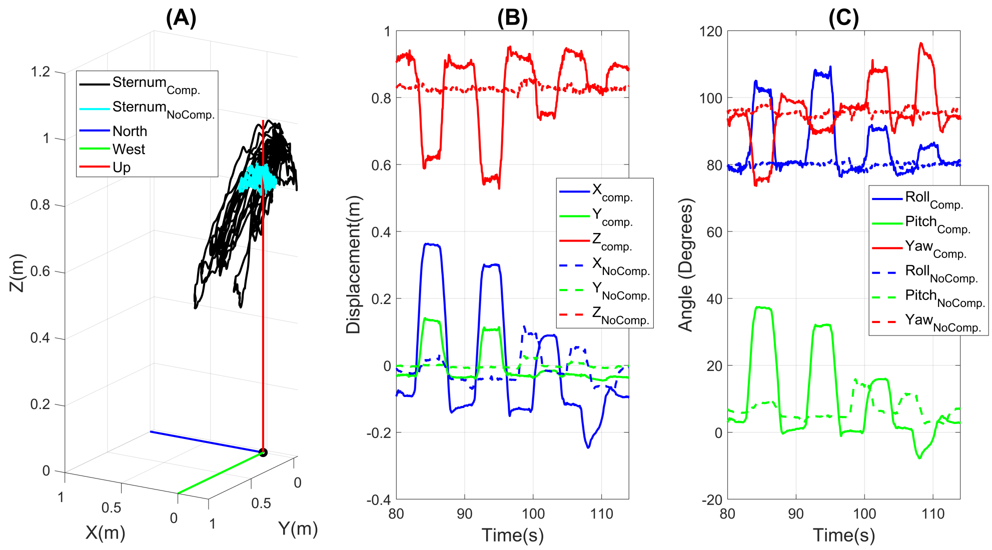

Figure 14. Displacement and rotation of trunk during normal and extended reaching tasks on table. This figure is intuitive in terms of distinction of tasks including compensatory movements from those without compensation strategies. (A) 3D displacement of sternum with and without compensatory movements, (B) Detailed displacement of sternum within tasks in-/excluding compensatory strategies at trunk level, (C) X-Y-Z Euler angles representation of trunk rotation during reaching tasks in the presence and absence of compensation. In figures (B) and (C), one cycle of reaching comprising reaching toward targets 2 to 5 as introduced in figure 9 is illustrated. Comp: compensation.

4. Discussion

In this study, we developed a HAR system capable of identifying reaching events from the other activities of daily living. Amongst the detected reaching events, our HAR system can also distinguish planar reaching events from the reaching events performed in 3D space. Up to now, most of the literature with the aim of quality of upper-limb movement assessment solely investigated the quality of 2d-planar reaching events. However, we firmly believe that within ADL there exists a combination of reaching tasks in both 3D space and transverse 2D planes. Furthermore, it is beneficial to classify 2D and 3D reaching tasks separately as the metrics for assessing the quality of movement might be different for each case. Deducing from the second stroke recovery and rehabilitation roundtable (SRRR), Kwakkel et al [46] recommend that dimensionality of performing tasks to assess the quality of movement should be a mixture of 2D and 3D rather than being 2D or 3D alone. They also proposed the list of suggested metrics to assess the quality of 2D and 3D movements separately, where some metrics such as smoothness or displacement overlap between both groups. Conversely, some metrics such as index of curvature (IC) are defined for 2D movements only, whereas others such as peak angular velocity of elbow extension are only defined for 3D movements [47]. Therefore, the HAR system proposed in this study simplifies the procedure of assessing the quality of 2D and 3D movements separately.

Each of the static hand poses in this study (see figure 8) can be related to a specific pose during ADL. Due to the lack of sensors on the fingers, it was not possible to inspect more functional hand postures in this study. The finger sensors were not added to our proposed system as they considerably diminish the comfortability of the system for ADL usage. But, kinetic and kinematic analysis of functional hand postures and grasps were shown to be useful to investigate stroke recovery and rehabilitation progress. Kwakkel et al [46] proposed that finger individuation and grip between thumb and index finger are necessary performance assays to be appraised. Therefore, in the future study working on stroke patients, it is of interest to also assess functional hand postures and grasps. Although wearing sensors on the fingers is remarkably impeding during ADL, one can define some specific exercises such as the ones similar to ARAT test to be done on a weekly basis and in home environment. So, finger sensors can be worn only during these weekly exercises. The iHAM system [59] is a minimal sensing setup in terms of the number of sensors on fingers, comprising one IMU on each of the thumb, index, and middle fingers in combination with pressure sensors on the fingertips. We propose the applicability of using the iHAM system during home-based weekly ARAT tests to both take care of our minimal sensing aim and take the functional hand postures and grasps assessment into consideration. This system was already validated on stroke patients and showed the feasibility of analyzing kinetics and kinematics [59].

The main classification problem of this work was subdivided into two sub-problems of classifying static and dynamic activities separately. We firstly tried to distinct between static and dynamic activities using an ANN; however, the confusion amongst static and dynamic classes led us to find an alternative. So, we substituted the ANN with a signal-conditioning approach called motion detection algorithm which was based on the fluctuations in the norm of acceleration signal. Then we fed the identified static and dynamic activities separately to the corresponding ANN-based classifiers. This way of redefining the main problem as two simpler sub-problems is beneficial due to simplifying the structure of ANN models, requiring less number of features, decreasing computational costs, and avoiding classification of a dynamic activity as a static given an "ideal" threshold of motion detection. However, our results showed the complexity of finding the threshold close to the "ideal" value.

Finding the optimum value as the threshold of motion detection is challenging during ADL and is prone to inaccuracy. There is indeed a trade-off for setting this threshold per subject, as too large values of threshold would lead to missing slower movements and conversely when the threshold is comparatively too low, some static poses can be misclassified as dynamic activities. The former case was manifested when comparing the reaching events with and without hand support, where the ones on hand support were mostly seen to be slower. The latter case was mostly seen during stance in still pose, where in most cases thresholds of 0.10 - 0.11 m/s^2 led to wrongly detecting some instances of this pose as dynamic. Incorrectly classifying some time instances of stance still pose as dynamic was caused by the presence of hand tremor or its unconscious movements during stance when compared with the hand unconscious movements during sitting. That is because within the sitting posture, hands are usually laid on a supporting surface such as the arm of a chair or table. But, during stance, hands have more freedom to move even unconsciously. In future studies, the proposed subject-specific threshold can be improved by being adaptive over time, instead of using a constant value which is measured during the calibration period. An alternative method of motion detection can be made by setting the threshold of detection based on the norm of angular velocity vector instead of acceleration. This measure calculates the magnitude of angular rotation of the hand and can be proposed as a technique to detect movement. It is also possible to utilize the combination of both acceleration and angular velocity based methods to detect movement. This way, mutual movement detections should remain unchanged and the ones detected by only one method should be changed into static. Since the combination method makes use of two resources of information, one can expect it to work more accurately. However, future studies should inspect all of the suggested methods of motion detection in details to see which one overcomes the rest.

Although there is no rule of thumb for choosing the best machine learning algorithm to create a HAR, we utilized artificial neural network (ANN) in this work due to different reasons. Firstly, ANN is an eager supervised machine learning method that is already validated to be used in pattern recognition [60] and classification [61] problems. Moreover, ANN benefits from non-parametric [62] modelling which is a meritorious advantage when compared with the complex parametric and statistical methods, e.g.: Hidden Markov models. Furthermore, having a diversity of training algorithms such as conjugate gradient, Bayesian regularization, and Levenberg-Marquardt backpropagation [63], ANN gave us the possibility to create our models using different configurations. Amongst these training algorithms, we found conjugate gradient and Levenberg-Marquardt as the best training algorithms in terms of classification outcome for the preliminary and 10-fold cross-validation approaches, respectively. Most of the existing ANN-based HAR developments such as in [35–37] made use of convolutional neural network (CNN) as a subsection of deep learning. We believe these studies suffer from high computational costs and complexity. Therefore, hypothesizing the applicability of a simpler ANN-based approach, we showed the feasibility of developing a HAR system using multi-layer perceptron (MLP). Our proposed MLP-ANN is consisted of three layers including input, single hidden layer, and output. Simplifying the classification models is important in our study as the system is dealing with a considerably large dataset during ADL and enormous computations might lead to crashing the whole system or making it tremendously slow. Future work can inspect the applicability of unsupervised machine learning and clustering methods of classification in order to compare the outcomes with our developed HAR.

One of the most challenging steps of designing ANN for pattern recognition is how to take care of over-fitting and over-training issues. Random permutation of features during training process, while having the same number of training samples per class could ensure the absence of over-fitting. However, in most cases, classification problems are faced with imbalanced classes where the number of samples per class is unequal. Also in our work, data were not distributed over classes by the same proportion. In terms of the number of samples per class, $C1$ (palms to the bottom) and $C10$ (other activities) were the dominant static and dynamic classes, respectively. Our investigation showed that approximately 76% of the whole static samples belonged to $C1$ and 42% of the whole dynamic samples were owned by $C10$. Under imbalanced circumstances, the class with the largest amount of samples during training phase has more chance to be chosen as the output class of the upcoming test data. This situation is called over-fitting of the classifier to the class with the largest amount of samples. In this work, to assure overcoming imbalanced classes issue, firstly the number of samples per class was calculated. The class with the minimum number of samples was selected as the reference and the same amount of samples as in this reference class have randomly opted from the rest of classes. This way, one can create balanced classes to be used during training phase. The changes in the accuracy of classification while using balanced and imbalanced classes were obvious, where the imbalanced classification led to about 6% less accurate outcomes.

Comparing the preliminary and 10-fold cross-validation approaches of HAR development, one can conclude the 10-fold cross validation approach had better performance on classifying the dataset. The overall accuracy of classification using the 10-fold cross-validation approach was found to be 91.61% which was higher than overall accuracy of 87.88% resulted from the preliminary approach. Furthermore, the mean of accuracy, precision, sensitivity, specificity, and F1-score of dynamic classes in 10-fold cross-validation approach increased by 0.88%, 9.35%, 6.11%, 0.33%, and 7.82%, respectively, when compared with the preliminary approach results. These measures were also enhanced for classification of static classes by comparable proportions, namely: 0.86%, 9.28%, 7.00%, 0.62%, and 10.40% increment in the mean of accuracy, precision, sensitivity, specificity, and F1-score, respectively.

F1-score is considerably a prominent statistical measure in our study to evaluate the performance of the developed HAR. This measure calculates the harmonic mean of precision and sensitivity to give a single value as an estimate on positive detection performance of the HAR. Robustness in positive detection per class is significantly important in this work, as an appropriate assessment of quality of movements requires proper true positive detection of each class, especially for reaching events. Resulting from tables 2 and 4, one can conclude that F1-scores of all classes except $C6$ and $C7$ were improved using the 10-fold cross-validation when compared with preliminary approach. Among dynamic classes, this improvement is more obvious in the F1-score of $C9$ (3D reaching events) where this measure increased from 48.42% in preliminary approach to 68.99% in 10-fold cross validation approach. Also within static classes, F1-score in $C5$ (palm to the left - Horizontal) had the largest improvement, growing from 55.33% in preliminary approach into 94.17% in 10-fold cross validation. Although the F1-score decreased for $C6$ and $C7$ in 10-fold cross-validation approach, these reductions were not very noticeable.

Position estimation using IMUs is prone to inaccuracy due to the cumulative drift caused by the integration of measured acceleration over time. Adding a supplementary sensing modality has already shown the potential to overcome the velocity and position estimation drift issues [64,65]. Roetenberg et al [65] showed the feasibility of estimating ambulatory relative positions using a magnetic sensing modality comprising 3 orthogonal coils in combination with IMUs. Their method accomplished position estimation with accuracy of about 5 mm, when compared to the optoelectronics. Furthermore, Kortier et al [64] proposed a novel approach in which they merged IMUs with permanent magnet to estimate relative hand position with respect to sternum. They utilized two IMUs, placing one on sternum as the reference and another IMU in addition to a permanent magnet on hand. Magnetic tracing in their work resulted in drift-free estimation of position even for long periods of measurement such as more than one minute. To provide estimations of position, an extended Kalman filter was developed to fuse IMU estimates with the magnetic updates. Their resulting estimates of position showed average RMS error of 19.7 ± 2.2 mm using the proposed system when compared with the optoelectronics. The same approach as in [64] is also applicable in our work as it is not in contrast with the minimal sensing term. Although Kortier et al [64] suggested to put the sensors on the hand, we believe putting magnets on the forearms close to wrists could be more advantageous during ADL. This is because placing magnets on hand as the end effector may hinder performing several ADL especially those that hand should interact with ferromagnetic metals, e.g: iron. In future work, despite replacing the Madgwick filter with a more robust Kalman filter including positions in the states vector, utilizing a model with biomechanical constraints can also improve position estimation.

Based on our reference system output, the "absolute" hand displacement in global earth frame during extended reaching tasks on HS and table was estimated to be 34.34 cm $\pm$ 9.11 cm and 38.55 cm $\pm$ 8.55 cm, respectively. However, referring to table 6, the "relative" displacement of hand with respect to sternum was estimated as 26.77 cm $\pm$ 10.41 cm and 28.15 cm $\pm$ 9.46 cm within extended reach on HS and table, respectively. The noticeable difference among the absolute and relative displacement of hand emerges the existence of compensatory movement at trunk level. Furthermore, large standard deviations in table 6 were caused by the neglect of participants to move their hands to the exact final or initial hand positions. The recorded video reference of each experiment also confirmed the presence of these neglections.

Smoothness analysis of movement can be done using various metrics; however, the choice of the best smoothness metric among the existing metrics is still challenging. To unite the framework of smoothness analysis, Balasubramanian et al [66] analyzed a large set of metrics in terms of validity, sensitivity, reliability, and practicality. They reported *SPARC* to be more meritorious than the other metrics considering the aforementioned criteria. Scheltinga et al [32] also proposed a set of metrics which are suitable for assessing the stroke recovery. However, despite the efforts of [32] and [66] on unifying the framework, literature is still assessing smoothness using sets of metrics known as the best but not just one metric. Therefore, also in our work, we utilized a set of metrics that were reported as the most useful ones adapted from [20,32,56]. Future work should inspect the smoothness of movement within the target group of stroke survivors by the same metrics as used in this research to find the most representative one. Besides, investigating smoothness of movement in stroke survivors with different levels of impairment is also advantageous as it manifests which of the metrics can interpret the impairment level in the most distinctive way. Eventually, the metric with the most discriminative power can be proposed as the most relevant smoothness metric in quality of upper-limb movement assessment within stroke survivors.

In table 7, we reported the smoothness of different reaching-included activities. We represented the smoothness values by descriptive statistics in the form of *mean* $\pm$ *standard deviation* for each of the metrics. However, analyzing descriptive statistics such as mean only interprets the existing dataset and does not let us investigate the generalizability of results. But, using inferential statistics one can use the available dataset to inspect the generalizability of results. In this work, we defined a research question regarding the generalizability of smoothness measurements as follows: "Does the measurement system, namely reference or minimal sensing setup create a significant difference between the output of each of the smoothness metrics within different tasks?". To answer this question, we utilized two-sample t-test with the confidence interval of 95%. Overall, the means of LDJ and $SPARC_{Acc}$ estimations were found to be significantly different between systems (p-value = 0.00). Contrarily, there was no significant difference between the estimated means of $SPARC_{Vel}$, CM, and NP (for details see table A1 in appendix). This shows that $SPARC_{Vel}$, CM, and NP were not considerably sensitive to the measurement system, when compared with LDJ and $SPARC_{Acc}$. Despite the equation of each metric that can be interpreted to emerge the sensitivity of output to the minor changes in input, there might be some other resources that affected the measurements. As a case in point, although sensors were rigidly attached on top of each other at the beginning of measurements, they might be slipped from their initial position during experiment. Furthermore, the minimal sensing system and full-body capturing system were using different generations of Xsens IMUs. So, the precision of sensors and consequently the output of measurements was expected not to be completely identical. Future study can investigate the main resources of emerging these differences in more details.

Correlation metric (CM) calculates the similarity of a given velocity profile with a minimum jerk velocity curve. Our study showed that even for the healthy participants performing fairly smooth movements, this metric cannot reach the maximum value. As represented in table 7, the largest mean value of CM was found to be approximately 84% for the extended reaching events on HS which is still 16% less correlated with an ideal minimum jerk profile. Also, other smooth movements such as normal reaching on HS and table and extended reaching on table were all less than 79% correlated with the reference minimum jerk profile. Furthermore, the large standard deviations of CM represent inconsistency of this metric over different trials of the same activity. Another drawback of CM is that this metric compares a given velocity curve with a symmetric profile. Despite reach-to-point events, reach-to-grasp events were seen to be mostly asymmetric movements in terms of their velocity profile. This emerges the non-applicability of using this metric on asymmetric movements such as reach-to-grasp. Thus, the idealism of the reference velocity curve [58] and symmetrical characteristics of CM can be referred to as limitations of this smoothness metric. To overcome these restrictions, Scheltinga et al [32] tried to replace the minimum jerk profile with a simulated asymmetric curve to be used as a reference for asymmetrical movements such as reach-to-grasp. They utilized reach-to-grasp velocity profiles of over 100 healthy participants to estimate the best fitting polynomial among them. This profile was created to be a reference for the assessment of reach-to-grasp movements of stroke survivors; however, applying this method also has its constraints. Since the estimated polynomial is created based on the data from one specific study, generalizability of the proposed polynomial to the other studies is not guaranteed. Furthermore, developing such a best fitting profile per study makes the outcomes of different studies incomparable due to having diverse references. Thus, we propose that correlation metric cannot be known as a representative smoothness metric on account of the mentioned limitations.

In this study, we proposed that smoothness analysis solely may not be a comprehensive metric of quality of movement assessment. We showed that there exist some quite smoothness movements in which the relative displacement of hand and trunk is relatively short. Examples of these kinds of movements are the extended reaching tasks on HS or table as a part of our experiments, where the subject tried to bend over the targets to reach a farther distance. Deriving from table 6, our measurements based on reference system showed that the relative displacement between hand and sternum slightly raised from 28.05 ± 8.47 *cm* into 28.15 ± 9.46 *cm* and slightly decreased from 26.84 ± 10.24 *cm* into 26.77 ± 10.41 *cm* during extended reaching when compared with normal reaching tasks on table and HS, respectively. Although the targets were placed approximately 10 *cm* farther in the extended version of the reaching task, we showed that the mean of relative displacement between hand and sternum changed only by 0.10 *cm* and 0.07 *cm* on table and HS, respectively. This small variation in the relative displacement of hand and sternum during extended reach events emerges the presence of trunk flexion toward targets as the compensatory strategy. Therefore, one can conclude merely inspecting smoothness can be distracting in some cases, particularly when compensatory movements also exist. Consequently, we believe investigation of smoothness in addition to the relative displacement of hand with respect to sternum is more intuitive to appraise the quality of upper-limb movement.

The most challenging step after this study is how to transfer the proposed technology into the target group of stroke patients, as the final goal. Achieving this aim requires two intermediate steps. As the first step, an extended version of the current study needs to be done again on healthy subjects but in a real home environment and without any restrictions on the inclusion criteria of performing tasks. Thereafter, HAR performance can be evaluated in a more critical way and within a longer duration. During the extended experiments, one can record video as the reference for validating the performance of HAR within real ADL using GoPro (GoPro Inc, San Mateo, Calif) attached to the participants' chest or head. The second intermediate step is to validate the usage of IMUs for smoothness analysis. Although previous works have shown the feasibility of utilizing IMUs for kinematic analysis during ADL [3,15–17], literature still lacks a specific study to validate smoothness analysis derived from IMUs in comparison with optoelectronics.

The methods in this work were adapted to the healthy participants; however, dealing with stroke patients requires modifications at some points. As a case in point, the sensor to segment calibration in this study was done in N-pose, although this pose might be difficult to be held for a few seconds by patients. Fortunately, our calibration method can be done in different positions while facing north, e.g: sedentary. Moreover, the motion detection algorithm we defined in this study works based on the threshold of acceleration, although some stroke patients especially those with severe movement difficulties might not reach the proposed value of acceleration threshold within their movements. The proper threshold for the target group can be defined by investigating acceleration data from patients and finding the trade-off of defining a value which is neither too low nor too high. Other possibilities of motion detection can also be tested, as earlier discussed in the fourth paragraph of the discussion section. Furthermore, to develop the HAR system, data were segmentized into windows of 1 second. This choice was made as healthy participants could accomplish lots of tasks such as reaching within a second, whereas smaller window sizes were not preferable on account of additional computational costs. Nevertheless, it is possible to use larger window sizes for the patients as their movements are not expected to be as quick as healthy participants. This can be done to create a system with more affordable computational costs. Moreover, in our experiments, we defined workspace determination task as the circular movements of the hand to estimate the "potential" reachable area of hand, but this task might be difficult to be done by patients. An alternative approach for determining patients' hand workspace is to estimate workspace during functional tasks, similar to the method in [19]. Based on our experimental protocol, one can draw an envelope that covers hand trajectory during reaching tasks in different directions, so, to use the reaching events also to estimate the hand workspace. This way, the term "potential" workspace can be replaced by the "functional" workspace of hand which is quite intuitive in ADL.

5. Conclusions

In this study, we developed a wearable system comprising 3 IMUs to assess the quality of upper-limb movements during ADL. The proposed system is capable of recognizing reaching events within ADL and subsequently assessing the quality of performing 2D-planar reaching tasks. To the best of our knowledge, this is the first HAR which is proposed to detect 2D-planar and 3D reaching events separately. This separation was found to be considerably advantageous as the outcome of the second stroke recovery and rehabilitation roundtable, since the metrics to assess 2D and 3D tasks are not necessarily identical. Our statistical analysis on the output of classifiers confirmed that the proposed HAR can reliably detect 2D reaching events, although the sensitivity of 3D reaching detection can be improved in future work. This improvement can be done by adding an altimeter sensor to the IMU placed on the forearm to allow having a better estimation of hand elevation. Moreover, our proposed system showed the feasibility of estimating validated smoothness metrics. Comparing descriptive statistics of smoothness analysis, our proposed system resulted in similar values to the ones calculated while using full-body suit. To investigate the generalizability of smoothness results, inferential statistics in the form of two-sample t-test was also applied. Overall, two-sample t-test indicated statistically significant differences in the outcomes of $SPARC_{Acc}$ and LDJ, resulted from the different sensing systems of the study. Nevertheless, $SPARC_{Vel}$, NP, and CM were seen not to be considerably dependent on the measurement system. Despite smoothness, the relative displacement between hand and sternum in addition to the potential reachable workspace area were also estimated by the minimal sensing system and compared with the ground truth. The large error of displacement estimation in the results showed the necessity of adding an additional sensing modality in future work, e.g: permanent magnets to estimate drift-free displacements by magnetic tracing. Eventually, although the current study investigated the performance of the proposed system on six healthy subjects, the final aim is to transfer this technology to the target group of stroke patients. To do this, two intermediate studies should be done, namely: an extended version of the current study to be done in a real home environment, and a study which aims at the validation of smoothness analysis using IMUs when compared with optoelectronics. Afterward, the proposed system is ready to be tested on stroke patients.

Abbreviations

The following abbreviations are used in this manuscript:

Acc	Acceleration
ADL	Activities of daily living
ANN	Artificial neural network
ARAT	Action research arm test
BLCS	Body local coordinate system
CNN	Convolutional neural network
DWT	Discrete wavelet transform
FFT	Fast Fourier transform
FMA-UE	Fugl-Meyer assessment for upper extremity
HAR	Human activity recognizer
HMM	Hidden Markov model
HS	Hand support
IMU	Inertial measurement unit
LDJ	Logarithm of dimensionless jerk
LS-SVM	Least square support vector machine
MLP	Multi-layer perceptron
MRMR	Maximum relevance minimum redundancy
NP	Number of peaks
Ref.	Reference
SPARC	Spectral arc length
SVM	Support vector machine
Vel	Velocity

Appendix A Stroke

According to World Health Organisation (WHO), cardiovascular diseases (CVDs) are the main cause of death worldwide and around 17.9 million deaths accounting for 31% of universal deaths in 2016 were caused by CVDs [67]. Within these deaths, approximately 85% of them caused by a heart attack and stroke [67]. Stroke has two main types, namely: ischemic and hemorrhagic. The former is caused by the existence of blood clot called *Thrombus* that narrows or completely blocks the artery and hinders normal blood flow. Subsequently, lack of blood flow and oxygen level in the following brain tissues cause them to become starved. But, hemorrhagic stroke is caused by abnormal bursting a blood vessel such that blood directly spills into brain tissues. The incidence of ischemic type is found to be quite higher than hemorrhagic. Unfortunately, the pre-stroke well-being conditions cannot be recovered completely in stroke survivors due to the impossibility of fully retrieving the functionality of the damaged tissues of the brain.

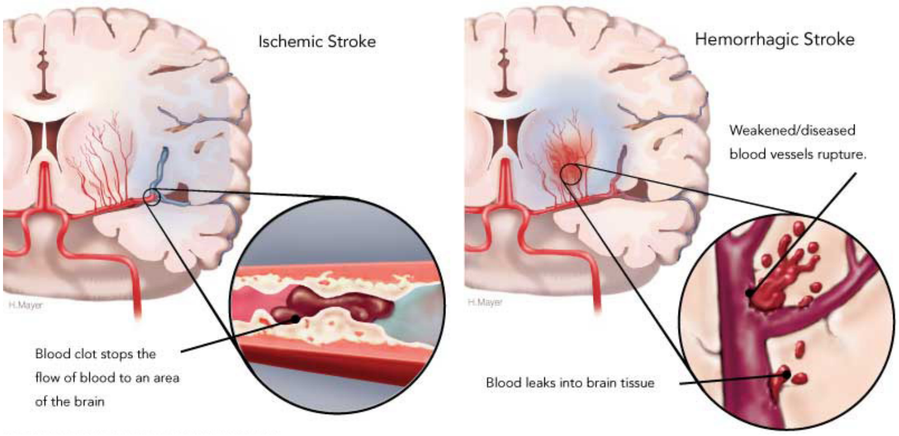

Figure A1. Two main types of stroke and their causes, adapted from [68].

In addition to high mortality rate, stroke is the second most common cause of disability throughout the world [1] and about 80% of stroke patients suffer from Hemiparesis [69]. Hemiparesis is the condition when the person has infirmity or is unable to move one side of the body. Moreover, almost half of the stroke patients endure long-term upper-limb motor dysfunctionalities [2] which affect their performance during ADL. It is shown that rehabilitation has a positive influence on regaining the movement capabilities [69], although it is also quite prominent to differentiate between restitution and substitution of movement patterns [5]. Various treatments exist to enhance the movement abilities of the affected side of stroke survivors comprising the following [69]:

- *Modified constraint-induced therapy (mCIT):* Restricting movements by the contralateral side such that affected side becomes more active.
- *Electrical stimulation:* Placement of electrical pads on the weakened muscles to help them perform contraction.
- *Cortical stimulation:* Placing electrode on the dura that emits an electrical signal to the cortex while performing rehabilitation exercises.
- *Mental Imagery:* The nerves in the brain contributing to physical movements and Imagination overlap such that visualization of performing tasks by the affected side can be effective when combined with other treatments.
- *Assistive Devices:* using supports such as walkers or wheelchair based on physical therapist recommendation.

Appendix B Confusion matrices of preliminary approach of HAR development

Individual confusion matrix of each subject correspondent to the first approach of creating HAR is indicated in figure A2.

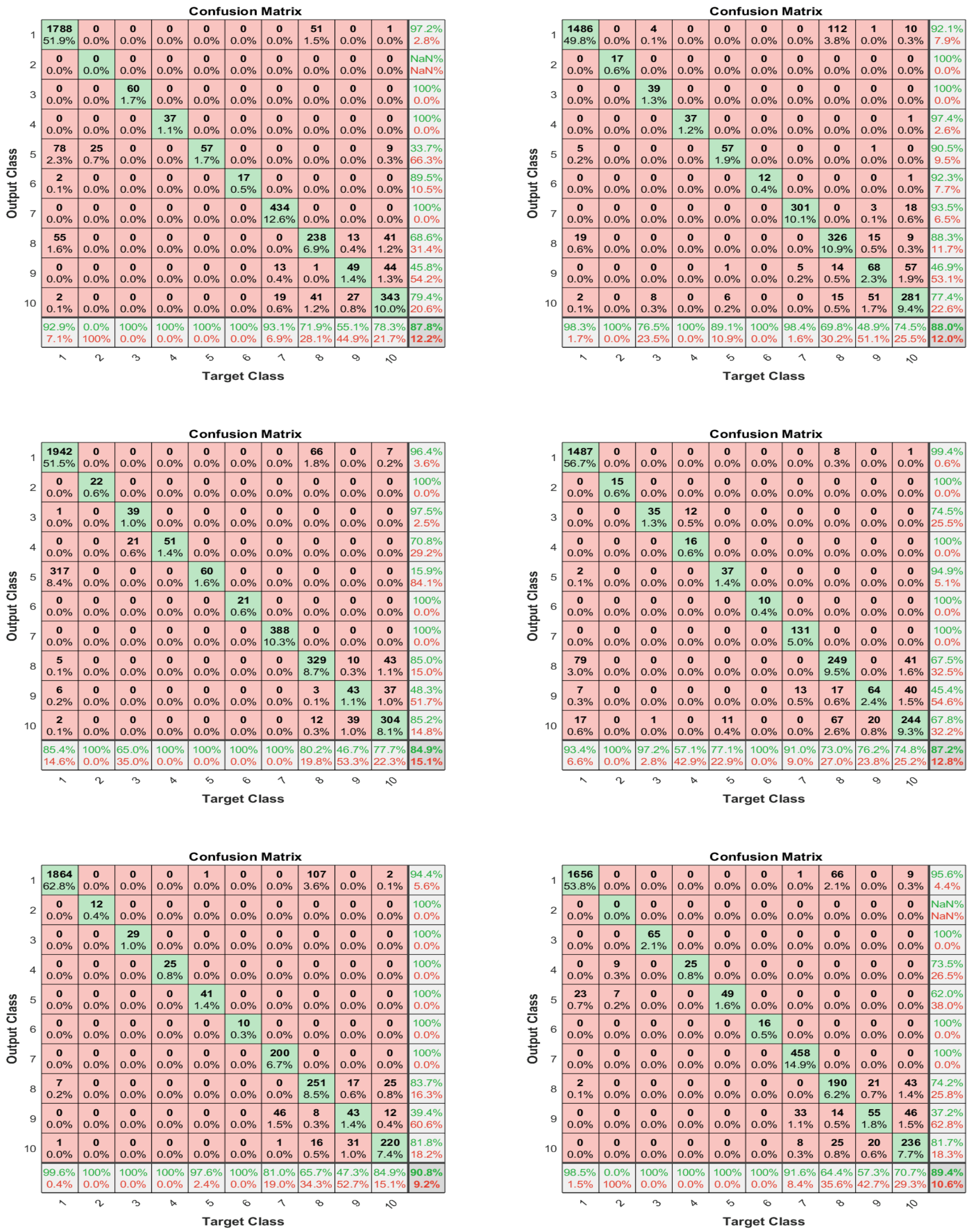

Figure A2. Confusion matrix of applying the first approach of HAR on each of the individuals. Correctly classified observations are shown in diagonal cells and incorrect predictions came in off-diagonal cells. Far right column of each confusion matrix depicts precision and false discovery rate by green and red, respectively. Far bottom row of each matrix shows recall and false negative rate by green and red, respectively. Overall accuracy comes in the bottom-right corner cell per matrix. C_i $(i \in 1, .., 10)$ stands for classes as introduced in table 1.

Appendix C Consistency of HAR performance over subjects

Due to the separability of static and dynamic ANN classifiers, one can inspect their performance individually. Since the most challenging part of classification problem is with the dynamic classes, figure A3 visualizes subject-specific analysis of statistical measures per dynamic class. This bar plot emerges the consistency of dynamic activities classification within subjects. It is not possible to investigate consistency per subject during cross-validation based method of HAR development, as the data needed to be shuffled. Therefore, consistency inspection was only done for the first approach of HAR creation, where model was created made based on researcher performance and tested on the participants.

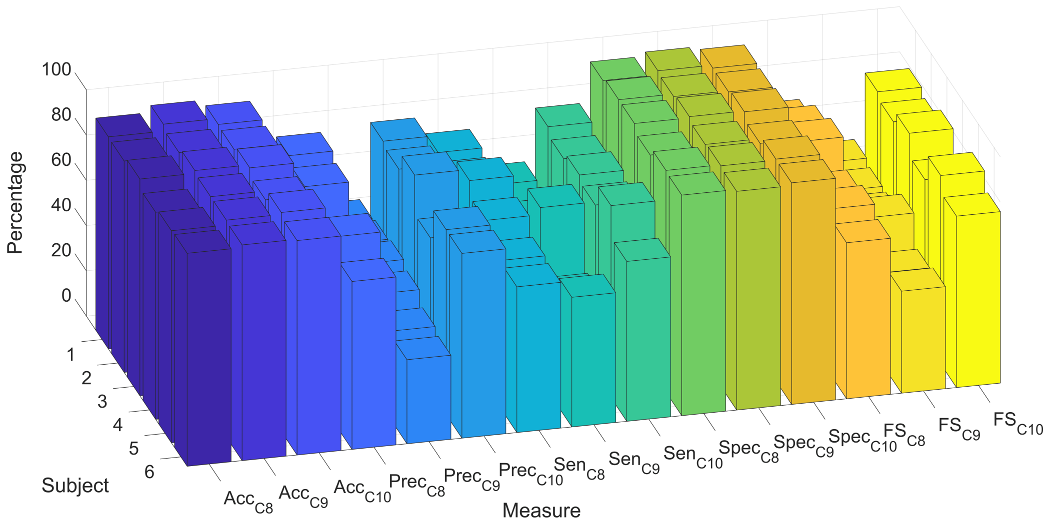

Figure A3. Subject-specific statistical analysis of HAR performance on dynamic classes. X-axis shows statistical metrics, namely: Acc: accuracy, $Prec$: precision, Sen: Sensitivity, $Spec$: specificity, FS: F1-score. C_i $(i \in \{8, 9, 10\})$ stand for classes, C8: 2D-planar reach, C9: 3d-reach, C10: other activities.

Appendix D Confusion matrices of 10-fold cross-validation approach of HAR development

In this section, the confusion matrices that were made during the 10-fold cross-validation approach of HAR development are shown. Applying 10-fold cross-validation led to having 10 models for each of dynamic and static classifiers. So, in total 20 confusion matrices comprising 10 for dynamic and 10 for static classification problems were made. To make it more readable, these 20 confusion matrices are depicted in 5 different figures, from figure A4 to A8, each comprising 4 confusion matrices.

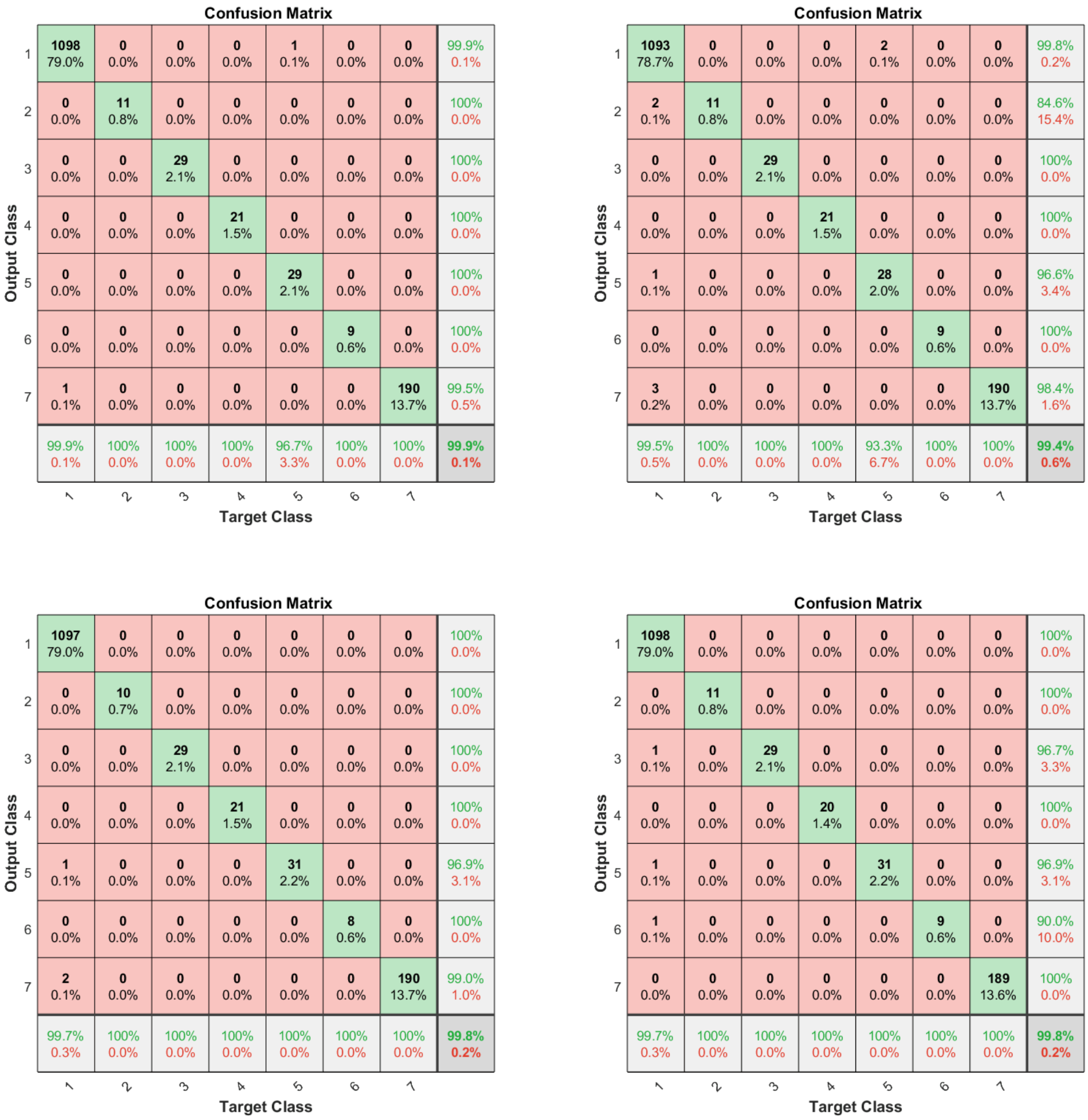

Figure A4. Confusion matrices of 10-fold cross-validation approach of HAR development (part 1/5). These matrices belong to the first 4 iterations of static models creation. C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose).

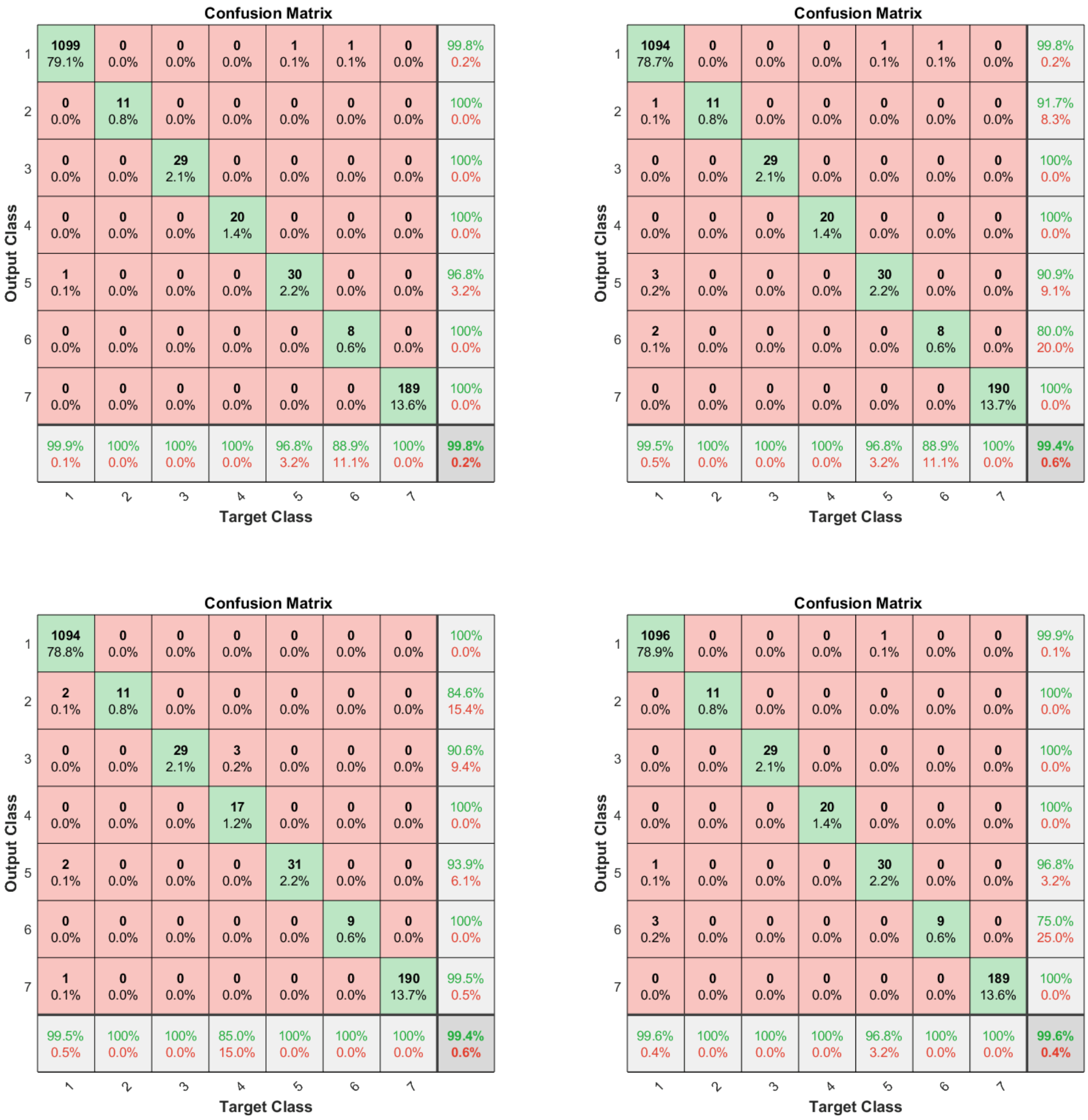

Figure A5. Confusion matrices of 10-fold cross-validation approach of HAR development (part 2/5). These matrices belong to the iterations 5 to 8 of static models creation. C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose).

Figure A6. Confusion matrices of 10-fold cross-validation approach of HAR development (part 3/5). The first row belongs to iterations 9 and 10 of static models creation and second row belongs to the first two iterations of dynamic models creation. In static models (first row): C1: Palm to bottom, C2: Palm to top, C3: Palm to left/front (Vertical), C4: Palm to right/person (Vertical), C5:Palm to left/person (Horizontal), C6: Palm to right (Horizontal), C7: Palms facing body (as in normal stance pose). In dynamic models, classes are defined as follows: C1: 2D-planar reach, C2: 3D reach, C3: Other activities.

Figure A7. Confusion matrices of 10-fold cross-validation approach of HAR development (part 4/5). These matrices belong to the iterations 3 to 6 of dynamic models creation. C1: 2D-planar reach, C2: 3D reach, C3: Other activities.

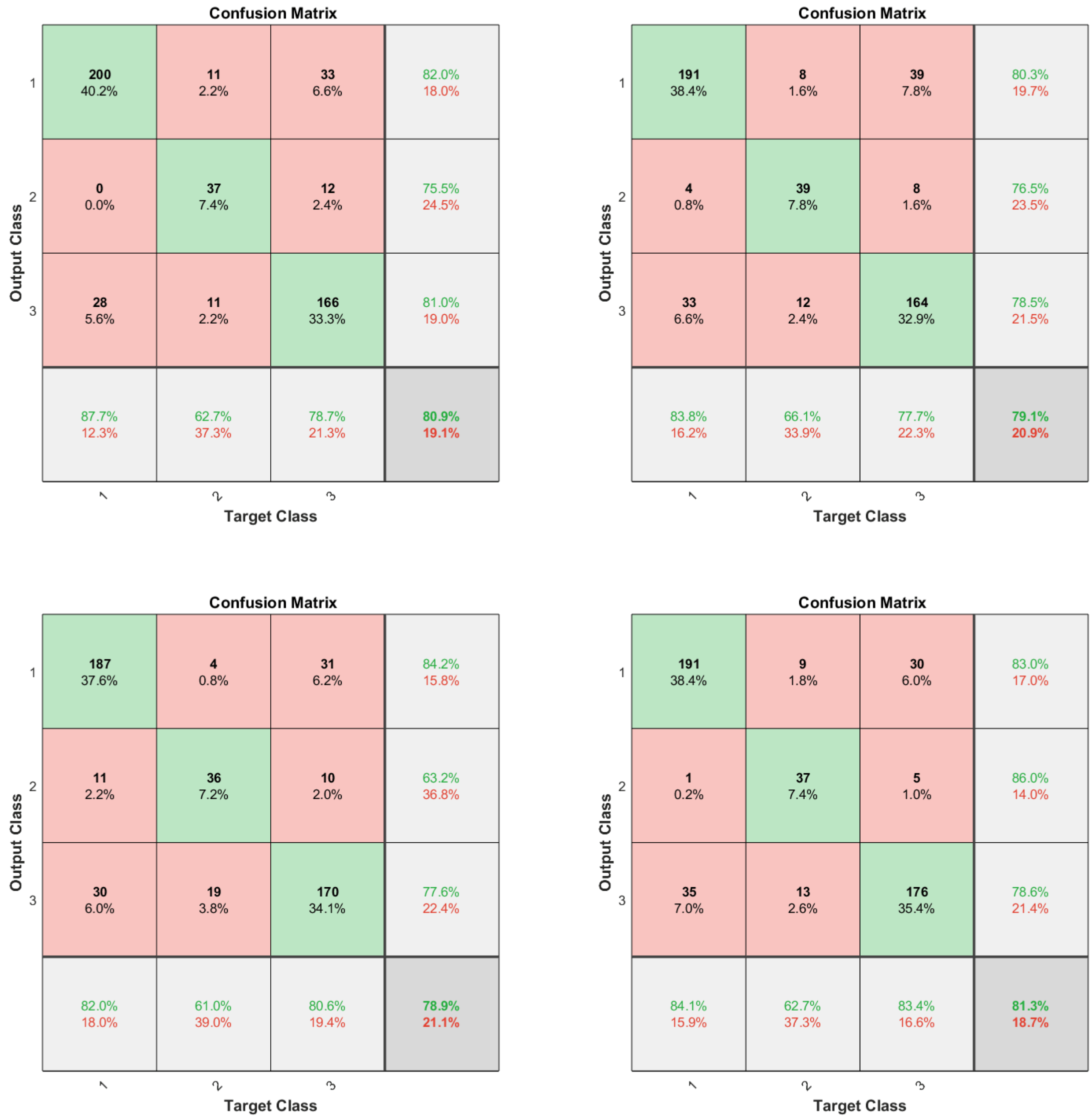

Figure A8. Confusion matrices of 10-fold cross-validation approach of HAR development (part 5/5). These matrices belong to the iterations 7 to 10 of dynamic models creation. C1: 2D-planar reach, C2: 3D reach, C3: Other activities.

Appendix E Consistency of displacement estimation

One can determine the distribution of displacement estimation error while doing the same task but over different trials. This shows the consistency of displacement estimation while performing reaching in different directions. As an illustration, figure A9 shows the error of estimated distance by the proposed system with respect to the MVN suit during performing reaching events on table. Extension and flexion portions of reaching tasks were also distinguished to investigate any probable difference.

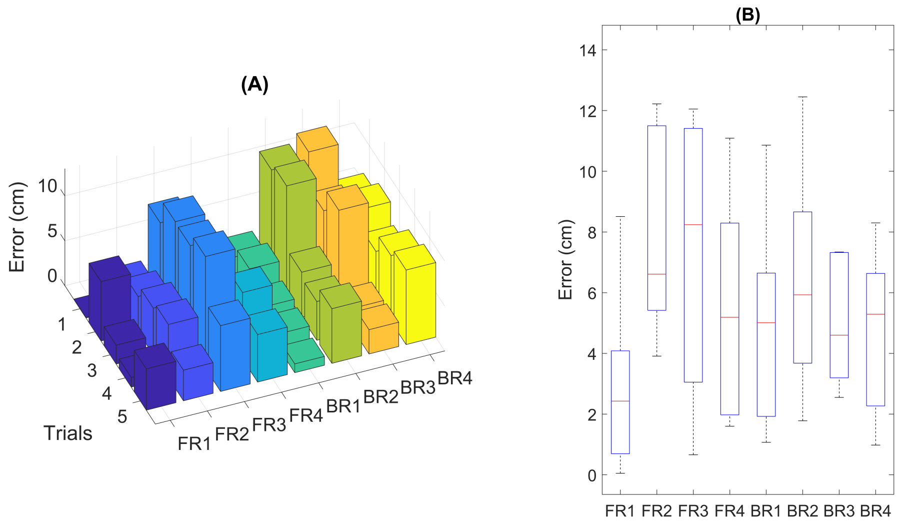

Figure A9. Error propagation over trials during displacement estimation of a sample participant while performing normal reach on table task. FR: forward reaching (Extension of hand), BR: Backward reaching (Flexion to the initial position). (A) Error of displacement estimation, where FR/BR 1 to 4, belong to the reaching to targets 2 to 5 of figure 9, respectively. (B) Corresponding boxplot of displacement estimation on (A) per reaching direction (forward/backward).

Appendix F Detailed results of two-sample t-test

The research question to be answered by the two-sample t-test is defined as follows: *Does the measurement system, namely full-body suit or minimal sensing setup create a significant difference between the output of each of the smoothness metrics within different tasks?*. So, the null hypothesis in here is: there is no statistically significant difference between the mean of each of the metrics measured by the reference versus the minimal sensing systems. Table A1 provides a detailed analysis of two-sample t-test resulted from SPSS software. The confidence interval was set to be 95%.

Table A1. Detailed results of applying two-sample t-test on the measured smoothness of various metrics within different reaching-included tasks of the study. Third and 4th columns (F-value and Sig.) interpret Levene's test for equality of variances and last three columns belong to the two-sample test for equality of means. The non-significant differences between the measurement systems are bolded. Sig.: significance level or p-value.
*: Not applicable. Since reach-to-grasp is not a symmetric movement, it is not comparable with a minimal jerk profile.

Task		Levene's test		Two-sample t-test		
	Metric	**F-value**	**Sig.**	**t**	**df**	**Sig. (2-tailed)**
Normal reach - HS	$SPARC_{Vel}$	18.35	0.00	-1.60	232.00	**0.11**
	$SPARC_{Acc}$	4.00	0.05	-5.05	310.62	0.00
	LDJ	7.05	0.01	-2.90	302.38	0.00
	CM	2.48	0.12	-0.11	297.00	**0.91**
	NP	0.01	0.95	-0.31	297.00	**0.76**
Extended reach - HS	$SPARC_{Vel}$	10.50	0.00	0.09	200.68	**0.92**
	$SPARC_{Acc}$	0.27	0.60	-3.25	295	0.00
	LDJ	4.27	0.04	-2.30	272.24	0.02
	CM	34.00	0.00	3.70	227.60	0.00
	NP	0.04	0.85	-0.30	284	**0.77**
Normal reach - table	$SPARC_{Vel}$	35.68	0.00	-2.91	277.71	0.00
	$SPARC_{Acc}$	10.89	0.00	-3.03	357.93	0.00
	LDJ	8.10	0.00	-3.66	363.66	0.00
	CM	0.41	0.52	-3.82	302	0.00
	NP	6.03	0.01	-1.89	372.28	**0.06**
Extended reach - table	$SPARC_{Vel}$	17.15	0.00	-1.92	278.85	**0.06**
	$SPARC_{Acc}$	4.40	0.04	-2.16	375.60	0.03
	LDJ	1.40	0.24	-3.94	392	0.00
	CM	18.87	0.00	1.32	360.20	**0.18**
	NP	3.76	0.05	-1.92	409	**0.06**
Reach-to-grasp	$SPARC_{Vel}$	8.22	0.01	-1.72	59.79	**0.09**
	$SPARC_{Acc}$	1.59	0.21	-1.53	117.00	**0.13**
	LDJ	8.09	0.01	-1.96	74.15	**0.06**
	CM^*	-	-	-	-	-
	NP	6.84	0.01	1.30	93.49	**0.20**
Overall	$SPARC_{Vel}$	28.18	0.00	-1.86	1240.98	**0.06**
	$SPARC_{Acc}$	17.54	0.00	-6.25	1437.14	0.00
	LDJ	23.81	0.00	-6.43	1400.17	0.00
	CM	20.89	0.00	0.55	1223.05	**0.59**
	NP	3.51	0.06	-1.91	1522.00	**0.06**

References

1. Murray, C.J.; Vos, T.; Lozano, R.; Naghavi, M.; Flaxman, A.D.; Michaud, C.; Ezzati, M.; Shibuya, K.; Salomon, J.A.; Abdalla, S.; others. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet* **2012**, *380*, 2197–2223.
2. Kwakkel, G.; Kollen, B.J.; van der Grond, J.; Prevo, A.J. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. *Stroke* **2003**, *34*, 2181–2186.
3. Van Meulen, F.B.; Klaassen, B.; Held, J.; Reenalda, J.; Buurke, J.H.; Van Beijnum, B.J.F.; Luft, A.; Veltink, P.H. Objective evaluation of the quality of movement in daily life after stroke. *Frontiers in bioengineering and biotechnology* **2016**, *3*, 210.
4. Xu, J.Y.; Wang, Y.; Barrett, M.; Dobkin, B.; Pottie, G.J.; Kaiser, W.J. Personalized multilayer daily life profiling through context enabled activity classification and motion reconstruction: An integrated system approach. *IEEE journal of biomedical and health informatics* **2014**, *20*, 177–188.
5. Kwakkel, G.; Kollen, B.; Lindeman, E. Understanding the pattern of functional recovery after stroke: facts and theories. *Restorative neurology and neuroscience* **2004**, *22*, 281–299.
6. Krakauer, J.W.; Carmichael, S.T.; Corbett, D.; Wittenberg, G.F. Getting neurorehabilitation right: what can be learned from animal models? *Neurorehabilitation and neural repair* **2012**, *26*, 923–931.
7. Levin, M.F.; Kleim, J.A.; Wolf, S.L. What do motor "recovery" and "compensation" mean in patients following stroke? *Neurorehabilitation and neural repair* **2009**, *23*, 313–319.
8. Alankus, G.; Kelleher, C. Reducing compensatory motions in video games for stroke rehabilitation. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2012, pp. 2049–2058.
9. Donkervoort, M.; Dekker, J.; Stehmann-Saris, F.C.; Deelman, B.G. Efficacy of strategy training in left hemisphere stroke patients with apraxia: A randomised clinical trial. *Neuropsychological rehabilitation* **2001**, *11*, 549–566.
10. Lai, S.M.; Studenski, S.; Duncan, P.W.; Perera, S. Persisting consequences of stroke measured by the Stroke Impact Scale. *Stroke* **2002**, *33*, 1840–1844.
11. Cirstea, M.; Levin, M.F. Compensatory strategies for reaching in stroke. *Brain* **2000**, *123*, 940–953.
12. Fugl-Meyer, A.R.; Jääskö, L.; Leyman, I.; Olsson, S.; Steglind, S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian journal of rehabilitation medicine* **1975**, *7*, 13–31.
13. Lyle, R.C. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. *International journal of rehabilitation research* **1981**, *4*, 483–492.
14. de los Reyes-Guzmán, A.; Dimbwadyo-Terrer, I.; Trincado-Alonso, F.; Monasterio-Huelin, F.; Torricelli, D.; Gil-Agudo, A. Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review. *Clinical Biomechanics* **2014**, *29*, 719–727.
15. Van Meulen, F.B.; van Beijnum, B.J.F.; Buurke, J.H.; Veltink, P.H. Assessment of lower arm movements using one inertial sensor. 2017 International Conference on Rehabilitation Robotics (ICORR). IEEE, 2017, pp. 1407–1412.
16. Veltink, P.H.; van Meulen, F.; van Beijnum, B.J.; Klaassen, B.; Hermens, H.J.; Droog, E.; Weusthof, M.H.; Lorussi, F.; Tognetti, A.; Reenalda, J.; others. Daily-life tele-monitoring of motor performance in stroke survivors. International Symposium: 3D Analysis of Human Movement 3D-AHM2014. Ecole Polytechnique Federale de Lausanne, 2014, pp. 159–162.
17. Patel, S.; Park, H.; Bonato, P.; Chan, L.; Rodgers, M. A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation* **2012**, *9*, 21.
18. Held, J.P.; Klaassen, B.; Eenhoorn, A.; Beijnum, B.J.F.v.; Buurke, J.H.; Veltink, P.H.; Luft, A.R. Inertial sensor measurements of upper-limb kinematics in stroke patients in clinic and home environment. *Frontiers in bioengineering and biotechnology* **2018**, *6*, 27.
19. van Meulen, F.B.; Reenalda, J.; Buurke, J.H.; Veltink, P.H. Assessment of daily-life reaching performance after stroke. *Annals of biomedical engineering* **2015**, *43*, 478–486.
20. Balasubramanian, S.; Melendez-Calderon, A.; Burdet, E. A robust and sensitive metric for quantifying movement smoothness. *IEEE transactions on biomedical engineering* **2011**, *59*, 2126–2136.

21. Hogan, N.; Sternad, D. Sensitivity of smoothness measures to movement duration, amplitude, and arrests. *Journal of motor behavior* **2009**, *41*, 529–534.
22. Milner, T.E. A model for the generation of movements requiring endpoint precision. *Neuroscience* **1992**, *49*, 487–496.
23. Krebs, H.I.; Hogan, N.; Aisen, M.L.; Volpe, B.T. Robot-aided neurorehabilitation. *IEEE transactions on rehabilitation engineering* **1998**, *6*, 75–87.
24. Rohrer, B.; Fasoli, S.; Krebs, H.I.; Hughes, R.; Volpe, B.; Frontera, W.R.; Stein, J.; Hogan, N. Movement smoothness changes during stroke recovery. *Journal of Neuroscience* **2002**, *22*, 8297–8304.
25. Romero, D.H.; Van Gemmert, A.W.; Adler, C.H.; Bekkering, H.; Stelmach, G.E. Altered aiming movements in Parkinson's disease patients and elderly adults as a function of delays in movement onset. *Experimental Brain Research* **2003**, *151*, 249–261.
26. Kim, N.H.; Wininger, M.; Craelius, W. Training grip control with a Fitts' paradigm: A pilot study in chronic stroke. *Journal of Hand Therapy* **2010**, *23*, 63–72.
27. Hermsdörfer, J.; Hentze, S.; Goldenberg, G. Spatial and kinematic features of apraxic movement depend on the mode of execution. *Neuropsychologia* **2006**, *44*, 1642–1652.
28. Gulde, P.; Hermsdörfer, J. Both hands at work: the effect of aging on upper-limb kinematics in a multi-step activity of daily living. *Experimental brain research* **2017**, *235*, 1337–1348.
29. Metrot, J.; Mottet, D.; Hauret, I.; van Dokkum, L.; Bonnin-Koang, H.Y.; Torre, K.; Laffont, I. Changes in bimanual coordination during the first 6 weeks after moderate hemiparetic stroke. *Neurorehabilitation and neural repair* **2013**, *27*, 251–259.
30. Adams, R.J.; Lichter, M.D.; Krepkovich, E.T.; Ellington, A.; White, M.; Diamond, P.T. Assessing upper extremity motor function in practice of virtual activities of daily living. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2014**, *23*, 287–296.
31. Baur, B.; Fürholzer, W.; Jasper, I.; Marquardt, C.; Hermsdörfer, J. Effects of modified pen grip and handwriting training on writer's cramp. *Archives of physical medicine and rehabilitation* **2009**, *90*, 867–875.
32. Scheltinga, B. Suitable metrics for upper-limb movement smoothness during stroke recovery. Master's thesis, University of Twente, 2019 (in progress).
33. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. International conference on pervasive computing. Springer, 2004, pp. 1–17.
34. Bonomi, A.G.; Goris, A.H.; Yin, B.; Westerterp, K.R. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine & Science in Sports & Exercise* **2009**, *41*, 1770–1777.
35. Ordóñez, F.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115.
36. Yang, J.; Nguyen, M.N.; San, P.P.; Li, X.L.; Krishnaswamy, S. Deep convolutional neural networks on multichannel time series for human activity recognition. Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.
37. Zeng, M.; Nguyen, L.T.; Yu, B.; Mengshoel, O.J.; Zhu, J.; Wu, P.; Zhang, J. Convolutional neural networks for human activity recognition using mobile sensors. 6th International Conference on Mobile Computing, Applications and Services. IEEE, 2014, pp. 197–205.
38. Pober, D.M.; Staudenmayer, J.; Raphael, C.; Freedson, P.S.; others. Development of novel techniques to classify physical activity mode using accelerometers. *Medicine and science in sports and exercise* **2006**, *38*, 1626.
39. He, J.; Li, H.; Tan, J. Real-time daily activity classification with wireless sensor networks using hidden markov model. 2007 29th Annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2007, pp. 3192–3195.
40. Su, S.; Wang, L.; Celler, B.; Ambikairajah, E.; Savkin, A. Estimation of walking energy expenditure by using support vector regression. 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. IEEE, 2006, pp. 3526–3529.
41. Mannini, A.; Intille, S.S.; Rosenberger, M.; Sabatini, A.M.; Haskell, W. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise* **2013**, *45*, 2193.
42. Zhang, S.; Rowlands, A.V.; Murray, P.; Hurst, T.L.; others. Physical activity classification using the GENEA wrist-worn accelerometer. PhD thesis, Lippincott Williams and Wilkins, 2012.

43. Liu, S.; Gao, R.X.; John, D.; Staudenmayer, J.W.; Freedson, P.S. Multisensor Data Fusion for Physical Activity Assessment. *IEEE Transactions on Biomedical Engineering* **2012**, *59*, 687–696. doi:10.1109/TBME.2011.2178070.
44. Wan, J.; O'grady, M.J.; O'hare, G.M. Dynamic sensor event segmentation for real-time activity recognition in a smart home context. *Personal and Ubiquitous Computing* **2015**, *19*, 287–301.
45. Zhan, Y.; Kuroda, T. Wearable sensor-based human activity recognition from environmental background sounds. *Journal of Ambient Intelligence and Humanized Computing* **2014**, *5*, 77–89.
46. Kwakkel, G.; Van Wegen, E.; Alt-Murphy, M.; Krakauer, J.; Levin, M.; Van Dokkum, L.; Winstein, C.; Burridge, J.; Melendez-Calderon, A.; Keller, T.; Kitago, T.; Nordin, N.; Pomeroy, V.; Veerbeek, J.; van Wijck, F. Standardized measurement of quality of upper limb movement after stroke: Consensus-based recommendations for the use of metrics from the 2nd Stroke Recovery and Rehabilitation Roundtable. (Pre-print).
47. Murphy, M.A.; Murphy, S.; Persson, H.C.; Bergström, U.B.; Sunnerhagen, K.S. Kinematic Analysis Using 3D Motion Capture of Drinking Task in People With and Without Upper-extremity Impairments. *JoVE (Journal of Visualized Experiments)* **2018**, p. e57228.
48. Madgwick, S. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. *Report x-io and University of Bristol (UK)* **2010**, *25*, 113–118.
49. Kohavi, R.; others. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai. Montreal, Canada, 1995, Vol. 14, pp. 1137–1145.
50. Zheng, Y. Human activity recognition based on the hierarchical feature selection and classification framework. *Journal of Electrical and Computer Engineering* **2015**, *2015*, 34.
51. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **2005**, pp. 1226–1238.
52. Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology* **2005**, *3*, 185–205.
53. Choe, N.; Zhao, H.; Qiu, S.; So, Y. A Sensor-to-Segment Calibration Method for Motion Capture system based on low cost mimu. *Measurement* **2019**, *131*, 490–500.
54. Bachmann, E.R.; Duman, I.; Usta, U.; McGhee, R.B.; Yun, X.; Zyda, M. Orientation tracking for humans and robots using inertial sensors. Proceedings 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation. CIRA'99 (Cat. No. 99EX375). IEEE, 1999, pp. 187–194.
55. Giansanti, D.; Maccioni, G.; Macellari, V. The development and test of a device for the reconstruction of 3-D position and orientation by means of a kinematic sensor assembly with rate gyroscopes and accelerometers. *IEEE transactions on biomedical engineering* **2005**, *52*, 1271–1277.
56. Beck, Y.; Herman, T.; Brozgol, M.; Giladi, N.; Mirelman, A.; Hausdorff, J.M. SPARC: a new approach to quantifying gait smoothness in patients with Parkinson's disease. *Journal of neuroengineering and rehabilitation* **2018**, *15*, 49.
57. Teo, C.L.; Burdet, E.; Lim, H. A robotic teacher of Chinese handwriting. Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002. IEEE, 2002, pp. 335–341.
58. Daly, J.J.; Hogan, N.; Perepezko, E.M.; Krebs, H.I.; Rogers, J.M.; Goyal, K.S.; Dohring, M.E.; Fredrickson, E.; Nethery, J.; Ruff, R.L. Response to upper-limb robotics and functional neuromuscular stimulation following stroke. *Journal of rehabilitation research & development* **2005**, *42*.
59. Claro Bhagubai, M. Quantification of pathological synergies in the upper extremity of stroke patients with the use of a multi-sensory system. Master's thesis, University of Twente, 2019 (in progress).
60. Bishop, C.M.; others. *Neural networks for pattern recognition*; Oxford university press, 1995.
61. Dreiseitl, S.; Ohno-Machado, L. Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics* **2002**, *35*, 352–359.
62. Tkacz, G.; others. *Non-parametric and Neutral Network Models of Inflation Changes*; Citeseer, 2000.
63. Tu, J.V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology* **1996**, *49*, 1225–1231.
64. Kortier, H.G.; Antonsson, J.; Schepers, H.M.; Gustafsson, F.; Veltink, P.H. Hand pose estimation by fusion of inertial and magnetic sensing aided by a permanent magnet. *IEEE transactions on neural systems and rehabilitation engineering* **2014**, *23*, 796–806.

65. Roetenberg, D.; Slycke, P.J.; Veltink, P.H. Ambulatory position and orientation tracking fusing magnetic and inertial sensing. *IEEE Transactions on Biomedical Engineering* **2007**, *54*, 883–890.
66. Balasubramanian, S.; Melendez-Calderon, A.; Roby-Brami, A.; Burdet, E. On the analysis of movement smoothness. *Journal of neuroengineering and rehabilitation* **2015**, *12*, 112.
67. World Health Organization. Cardiovascular diseases (CVDs). https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), 2017. Available online.
68. Heart and Stroke Foundation of Canada. What is stroke? https://www.heartandstroke.ca/stroke/what-is-stroke. Available online.
69. National Stroke Association. Hemiparesis. https://www.stroke.org/we-can-help/survivors/stroke-recovery/post-stroke-conditions/physical/hemiparesis/. Available online.