

Article

Assessing the Performance of Sensor Fusion Methods: Application to Magnetic-Inertial-Based Human Body Tracking

Gabriele Ligorio ^{1,*}, Elena Bergamini ², Iliaria Pasciuto ^{2,†}, Giuseppe Vannozzi ^{2,†}, Aurelio Cappozzo ² and Angelo Maria Sabatini ¹

¹ The BioRobotics Institute, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56125 Pisa, Italy; angelo.sabatini@sssup.it

² Department of Movement, Human and Health Sciences, Interuniversity Centre of Bioengineering of the Human Neuromusculoskeletal System, University of Rome "Foro Italico", Piazza Lauro de Bosis 15, 00135 Roma, Italy; elena.bergamini@uniroma4.it (E.B.); ilaria.pasciuto@uniroma4.it (I.P.); giuseppe.vannozzi@uniroma4.it (G.V.); aurelio.cappozzo@uniroma4.it (A.C.)

* Correspondence: g.ligorio@sssup.it; Tel.: +39-050-883-410; Fax: +39-050-883-101

† These authors contributed equally to this work.

Academic Editor: Simon X. Yang

Received: 22 November 2015; Accepted: 19 January 2016; Published: 26 January 2016

Abstract: Information from complementary and redundant sensors are often combined within sensor fusion algorithms to obtain a single accurate observation of the system at hand. However, measurements from each sensor are characterized by uncertainties. When multiple data are fused, it is often unclear how all these uncertainties interact and influence the overall performance of the sensor fusion algorithm. To address this issue, a benchmarking procedure is presented, where simulated and real data are combined in different scenarios in order to quantify how each sensor's uncertainties influence the accuracy of the final result. The proposed procedure was applied to the estimation of the pelvis orientation using a waist-worn magnetic-inertial measurement unit. Ground-truth data were obtained from a stereophotogrammetric system and used to obtain simulated data. Two Kalman-based sensor fusion algorithms were submitted to the proposed benchmarking procedure. For the considered application, gyroscope uncertainties proved to be the main error source in orientation estimation accuracy for both tested algorithms. Moreover, although different performances were obtained using simulated data, these differences became negligible when real data were considered. The outcome of this evaluation may be useful both to improve the design of new sensor fusion methods and to drive the algorithm tuning process.

Keywords: sensor fusion; algorithm benchmarking; inertial-magnetic sensors; human motion tracking; orientation; locomotion

1. Introduction

Sensor fusion is a signal processing technique that combines data measured by multiple sources in order to create a single measurement system with an augmented performance over each standalone sensor [1,2]. The reason for designing sensor fusion algorithms (SFAs) is two-fold: first, to improve the accuracy and/or robustness of the outcome by exploiting data redundancy and/or complementarity; second, to provide a complete picture of the phenomenon under investigation unifying the partial observations provided by each sensor.

SFAs are widely employed in several applications including autonomous navigation, robotics, environmental monitoring, and healthcare [2–5]. In particular, the sensor fusion based on observations

from magnetic and inertial sensors (commonly referred to as Magnetic-Inertial Measurement Units, MIMUs) is increasingly employed for the estimation of human body segment orientation in movement analysis and related applications [6–8]. A MIMU embeds a tri-axial gyroscope, accelerometer, and magnetic sensor that provide measurements of the body's angular velocity, the specific force (*i.e.*, the sum of the external and gravitational accelerations), and the Earth's magnetic field, respectively. The 3D body orientation can be estimated through the numerical integration of the rotational kinematic equation of a rigid body, by using angular velocity measurements and known initial conditions. The accelerometer and the magnetic sensors are supposed to track two external reference vectors: the Earth gravity acceleration (vertical reference vector) and the Earth magnetic field (heading reference vector). Estimating the body orientation by comparing the reference vectors measured in the body frame with their known counterparts expressed in an inertial reference frame, is known as field vector matching [9]. Usually, gyroscope integration and field vector matching are combined in an SFA to purposefully exploit their complementary properties. Several SFAs have been successfully proposed in the literature for the MIMU-based orientation estimation, which typically rely on either Kalman filtering [9–12] or complementary filtering [13,14]. On the contrary, when used alone, these sensors may yield poor results due to different issues characterizing the magnetic and the inertial sensors.

In this respect MIMU observations are disparate [3], in the sense that part of the orientation information is observed in three different physical domains, *i.e.*, the angular velocity, the specific force, and the Earth magnetic field vector. For this reason, MIMU data may be considered uncorrelated to each other, *i.e.*, no crosstalk among the three domains is expected. The only exception is the gyroscope sensitivity to the gravity and external accelerations [15] which, however, is often considered negligible as compared to other gyroscope error sources [16]. Within each domain, measurement noise is present in the sensor outputs [3], *i.e.*, each sample of measured angular velocity, specific force, and magnetic field exhibits a degree of uncertainty [17] which has different effects on the orientation estimates due to the different ways the MIMU sensor measurements are used to this purpose. Gyroscope-based tracking proved to be accurate during short-term rapid movements, although it is prone to boundless drift error over time occurring during integration. Of great concern in this regard is also the gyroscope bias, a slowly-varying output that is present even when the gyroscope is still. On the other hand, field vector matching does not suffer from drift errors, but it is heavily affected by external accelerations and magnetic disturbances, which are continuous and time-varying. Therefore, measurement noise, sensor bias, and external factors (external accelerations and ferromagnetic disturbances) result in conflicting information to be fused in the SFA, which may lead to highly inaccurate estimates of the 3D body orientation and even to the SFA failure to converge [3].

Unfortunately, when sensor observations are fused in an SFA, it is very difficult to assess to which extent each sensor issue influences the final error. This information would be crucial to guide the SFA design process (*i.e.*, the choice of different tuning settings or of the adaptive mechanisms to be built in the SFA) or to compare different combinations of sensor hardware components. Therefore, the development of benchmarking methods aimed at quantifying the effect of each sensor issue on the SFA performance is of the utmost importance. Nevertheless, no clear methodologies, guidelines, or tools for SFA performance assessment are available.

The main contribution of this paper is to propose a novel benchmarking method for the assessment of SFAs' performance. Another subsidiary contribution is to provide, as an outcome of the proposed benchmarking method, useful considerations about the MIMU-based orientation estimation research field. In fact, to show the potentialities of the proposed method and to illustrate its application, a case study is considered where the 3D orientation of a human body segment is estimated using a MIMU during a clinical test. The proposed benchmarking method relies on the concept of data hybridization: measured and simulated MIMU data are combined in different ways to create different testbeds, hereafter referred to as scenarios. Each scenario is conceived to isolate the influence of the errors characterizing each sensor and to allow quantifying the efficacy of adaptive mechanisms built in different SFAs. In order to highlight both these specific aspects of the benchmarking problem,

two different Kalman filter-based SFAs were considered and their performance in estimating the 3D orientation of the pelvis during a Timed Up and Go test was assessed.

2. Methods

2.1. Proposed Benchmarking Method: Overview

The proposed method for evaluating and comparing the performance of SFAs is based on the workflow depicted in Figure 1.

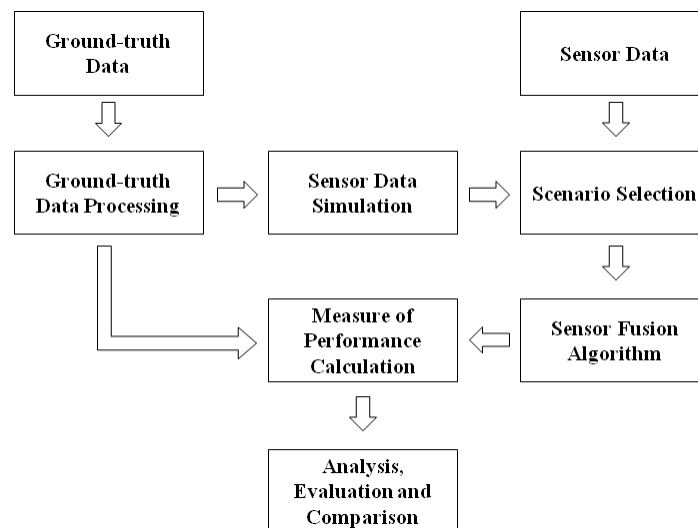


Figure 1. General overview of the proposed SFA evaluation framework.

Specifically, for the application investigated in the present study, ground-truth and real sensor data were acquired during a motion capture session. The ground-truth data were processed within the Ground-truth Data Processing block to obtain the reference output, which was used to simulate ideal sensor data (Data Simulation block). In addition it provided a reference for the calculation of the Measure Of Performance (MOP), *i.e.*, the metrics that quantifies the performance of an algorithm. A pool of evaluation scenarios (Scenario Selection block) was then created. Each scenario is a different combination of measured and/or simulated MIMU data. The SFA under analysis was then run using the data of each scenario as input, and the MOP was calculated by comparing the corresponding output with the reference one.

2.2. The Timed up and Go Dataset

The considered case study is the estimation of the pelvis orientation, by using data measured by a waist-worn MIMU, during a Timed Up and Go test (TUG) (Figure 2), which is a frequently performed clinical test aimed at assessing motor function. In particular, the TUG test was selected because it is a complex motor task, being characterized by transitory phases, such as sit-to-stand or change of direction, and by cyclic movements, such as level walking.

The University ethics board approved the adopted experimental protocol: 24 healthy subjects, after being informed of the goals and the modalities of the experiments, performed a TUG test. Prior to the TUG test, a MIMU (Opal, APDM Inc., Portland, OR, USA), embedding a tri-axial gyroscope, accelerometer, and magnetic sensor (± 6 g with $g = 9.81$ m/s², $\pm 1500^\circ$ /s and ± 600 μ T of full-range scale, respectively), was secured to the participants' lower back (lumbar region of the spine, approximately at L3–L4 vertebrae level, Figure 2), using an elastic belt. MIMU data were collected at 128 samples/s. A plastic plaque equipped (Figure 2) with a cluster of four infrared reflective markers was rigidly

attached to the MIMU case for ground-truth data acquisition using a nine-camera motion capture system (Vicon MX3, Oxford, UK) at 100 sample/s.

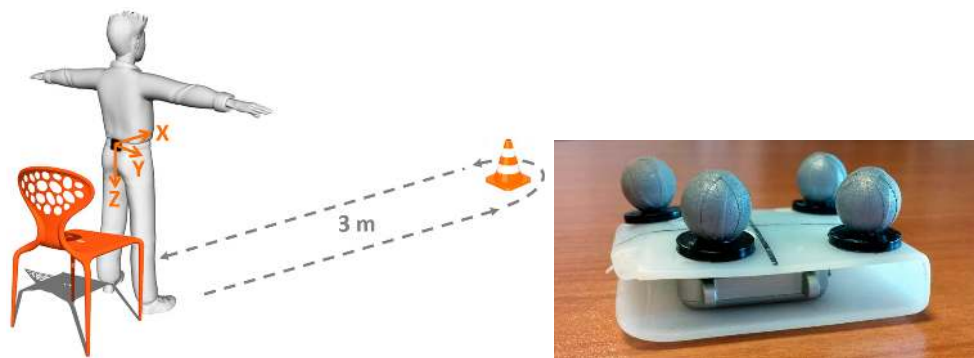


Figure 2. (Left) Sensor location on the participants' body, axes orientation and TUG scheme, and (Right) Opal MIMU with the infrared reflective marker cluster.

The reference frames used in this paper were the navigation frame $\{\mathbf{n}\}$ (also referred to as the global or inertial frame) and the body frame $\{\mathbf{b}\}$ (also referred to as the local frame). In the following, \mathbf{p}^b denotes the representation of the generic vector \mathbf{p} in $\{\mathbf{b}\}$, whereas \mathbf{q}^{bn} and \mathbf{R}^{bn} represent, respectively, the quaternion of orientation and the orientation matrix which rotate $\{\mathbf{n}\}$ to $\{\mathbf{b}\}$. Furthermore, $\hat{\mathbf{p}}$ is an estimate of the vector \mathbf{p} .

The MIMU sensor was calibrated before the experiments, as suggested in [6]. The navigation reference frame was defined with the z-axis aligned with the vertical direction, whereas the angle about the vertical axis was set as the value recorded during a static posture performed at the beginning of the trial. This static time window, which lasted 3 s, was also used to perform the gyroscope bias capture [6]. MIMU and stereophotogrammetric data streams were electronically synchronized using a square wave signal simultaneously detected by both systems. Both data were resampled at 200 Hz using cubic spline interpolation. Marker trajectories were low-pass filtered using a second-order zero-lag Butterworth filter. The cut-off frequency was determined by performing a residual analysis [18] on each trial of each subject and conservatively set to 6 Hz for all trials. A marker-cluster reference frame was then defined using the markers attached on the MIMU and its alignment with the MIMU case verified [19].

In order to obtain simulated (considered to be error-free and, thus, referred to as ideal) MIMU data, the ground-truth body orientation \mathbf{q}^{bn} and position \mathbf{b}^n , obtained from the marker cluster attached to the MIMU case, were used as an input to the data simulator described in [20]. In particular, for each trial, the ideal angular velocity was obtained by differentiating the ground-truth quaternions \mathbf{q}^{bn} with standard formulae [21]. The same quaternions were used to rotate the coordinates of the Earth's gravitational and magnetic field vectors, known *a priori*, from $\{\mathbf{n}\}$ to $\{\mathbf{b}\}$. The rotated Earth's magnetic field represented the ideal magnetometer data, whereas the external acceleration was added to the rotated gravitational field to build the ideal accelerometer data. All the ideal sensor data were corrupted with white noise before applying the stochastic filters (standard deviation estimated during the static postures equal to 5×10^{-3} rad/s, 5×10^{-3} m/s² and 0.15 μ T). At this point in time, both the experimental and the ideal data required to devise the different scenarios for the SFA evaluation were available.

2.3. Scenario Selection

To isolate the influence of the errors characterizing each MIMU sensor and to quantify the efficacy of the adaptive mechanisms built in to test SFAs, the following scenarios were defined (Table 1): (a) two scenarios where the full set of MIMU data was either simulated or measured (hereafter, they are referred to as the SIM and MEAS scenarios, respectively). These two scenarios represent the best and

worst testbeds for the considered SFAs, respectively; (b) three scenarios (named GYR, ACC, and MAG) in which two MIMU sensors' data were simulated and one MIMU sensor at a time was accounted for with its measured data. Specifically, the name of each scenario indicates the abbreviation of the sensor accounted for with its measured data. The rationale is to isolate the influence of each sensor data issue on the overall SFA output; (c) a scenario (named MOD) which differed from the SIM scenario in the way the simulated accelerometer data were generated: in this case, in fact, the simulated accelerometer data contained only the contribution due to the gravity reference vector (that determines the orientation estimation), without taking into account the external acceleration, which in this context represents a disturbance. In this way, the detrimental effect of the external acceleration on the SFAs performance was also evaluated.

Table 1. The six evaluation scenarios considered in this work, listed theoretically from the best to the worst case.

Scenario	Gyroscope	Accelerometer	Magnetometer
MOD	simulated	gravity-only	simulated
SIM	simulated	simulated	simulated
GYR	measured	simulated	simulated
ACC	simulated	measured	simulated
MAG	simulated	simulated	measured
MEAS	measured	measured	measured

2.4. Sensor Fusion Algorithms

Two Kalman-based orientation estimators were benchmarked in this work by using the proposed approach (Figure 3). The pseudo-code description of the two algorithms is reported in the Appendix A. In Algorithm 1, presented in [10], gyroscope measurements are used as input in two parallel linear Kalman filters that separate the component due to the reference field vectors (Earth's gravitational \mathbf{g}^b and magnetic \mathbf{h}^b fields) from the disturbances affecting the accelerometer and magnetometer readings (the external acceleration \mathbf{a}^b and the magnetic disturbances \mathbf{d}^b , respectively). The estimated reference field vectors $\hat{\mathbf{g}}^b$ and $\hat{\mathbf{h}}^b$ are then used to feed the TRIAD method [22] for single-frame orientation estimation, given that \mathbf{g}^n and \mathbf{h}^n are known.

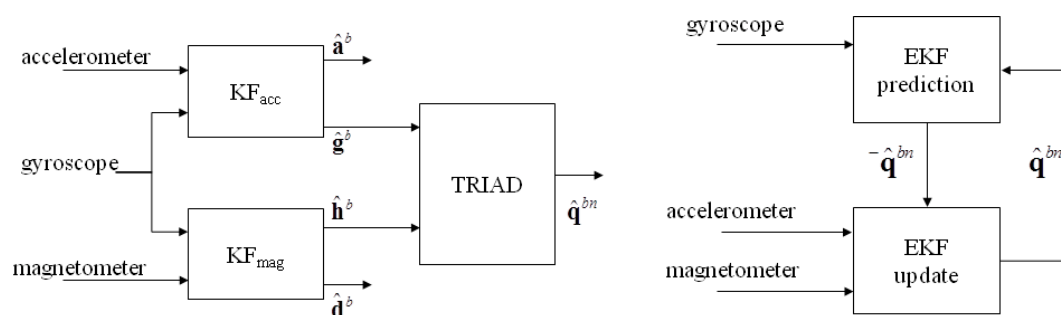


Figure 3. Overview of the Kalman-based SFAs considered in this work: (Left) Algorithm 1 and (Right) Algorithm 2.

Algorithm 2 is the Extended Kalman Filter (EKF) presented in [9]. In the prediction step, the quaternion estimate is projected ahead using the measured sample of the angular velocity. Then, assuming that external accelerations are negligible and no ferromagnetic disturbances are present, the measured acceleration and the measured magnetic field vector are used to update the projected quaternion in the nonlinear measurement equation. A linearization step is then needed at each iteration.

The combination of ideal and/or measured angular velocity, acceleration, and magnetic field vector data, associated with each above-mentioned scenario, were fed into both Algorithms 1 and 2.

For each SFA and time sample, six orientation estimates were obtained (one for each scenario) and each of them was compared to the ground-truth body orientation, as explained in Section 2.5.

2.5. Measure of Performance

The accuracy in estimating the MIMU orientation was expressed in terms of orientation error computed as the quaternion that rotates the estimated MIMU orientation $\hat{\mathbf{q}}_k^{bn}$ onto the ground-truth one \mathbf{q}_k^{bn} as proposed in [23]:

$$\Delta\hat{\mathbf{q}}_k^{bn} = \mathbf{q}_k^{bn} \otimes (\hat{\mathbf{q}}_k^{bn})^{-1} \quad (1)$$

where \otimes represents the quaternion multiplication operator and $(\mathbf{q})^{-1}$ is the inverse quaternion operator.

Then, in accordance with [10], $\Delta\hat{\mathbf{q}}_k^{bn}$ was split into two error components, one related to the heading $^{head}\Delta\hat{\mathbf{q}}_k^{bn}$ (yaw angle) and one to the attitude $^{att}\Delta\hat{\mathbf{q}}_k^{bn}$ (pitch and roll angles). The heading and attitude orientation parameterization, in fact, is particularly convenient for the problem at hand as the pitch and roll angles may be reasonably assumed to have similar error characteristics since they both represent an inclination with respect to the vertical direction [6,23]. The Root Mean Square (RMS) values of the scalar part (indicated as S in the following equations) of both $^{head}\Delta\hat{\mathbf{q}}_k^{bn}$ and $^{att}\Delta\hat{\mathbf{q}}_k^{bn}$ were then considered as MOPs:

$$\begin{aligned} \text{RMS}_{head} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (2\text{acos}[S(^{head}\Delta\hat{\mathbf{q}}_k^{bn})])^2} \\ \text{RMS}_{att} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (2\text{acos}[S(^{att}\Delta\hat{\mathbf{q}}_k^{bn})])^2} \end{aligned} \quad (2)$$

The MOP was computed for each scenario, SFA, and trial performed by each subject.

2.6. Statistical Analysis

In order to assess the performance of both Algorithms 1 and 2, the following specific questions were considered: (i) To which extent the SFA performance is affected by the errors characterizing each MIMU sensor? (ii) Can the SFA mitigate the effect of these sensor-specific errors, at least when all other sensors' data are considered to be as ideal? (iii) How do different SFAs behave when fed with either simulated or real sensor signals?

For what concerns questions (i) and (ii) (intra-algorithm analysis), a one way repeated-measures ANalysis Of VAriance (ANOVA) was performed, using the scenario as a factor, on both the attitude and heading accuracy obtained for each method, separately. The MOP values were transformed according to the Tukey Ladder of Powers in order to achieve normality, in case of a lack of normality revealed by the Shapiro-Wilk test. The Greenhouse-Geisser correction was used to take into account possible violations of the sphericity assumption. Given possible significant differences revealed by the ANOVA test for the scenario effect, the Dunn-Sidak's *post-hoc* pairwise tests were used to compare the GYR, ACC, and MAG scenarios with both the SIM and MEAS scenarios in order to answer (i) and (ii), respectively. In addition, the same test was adopted to compare the MOD and SIM scenario in order to assess the effect of the external acceleration on the SFAs output.

In order to answer question (iii) (inter-algorithm analysis), the presence of significant differences between the orientation accuracy obtained for Algorithms 1 and 2 was verified using a Wilcoxon test for both the SIM and MEAS scenarios. The alpha level of significance was set to 0.05 for all statistical tests.

3. Results

The heading and attitude ground-truth curves and the errors associated to each simulated scenario are reported in Figure 4 for one randomly-chosen participant. The different phases of the TUG task

are also indicated by different colors: sit-to-stand and stand-to-sit (yellow bands), walking (light blue bands), and turns of 180° about the cranio-caudal direction (green bands).

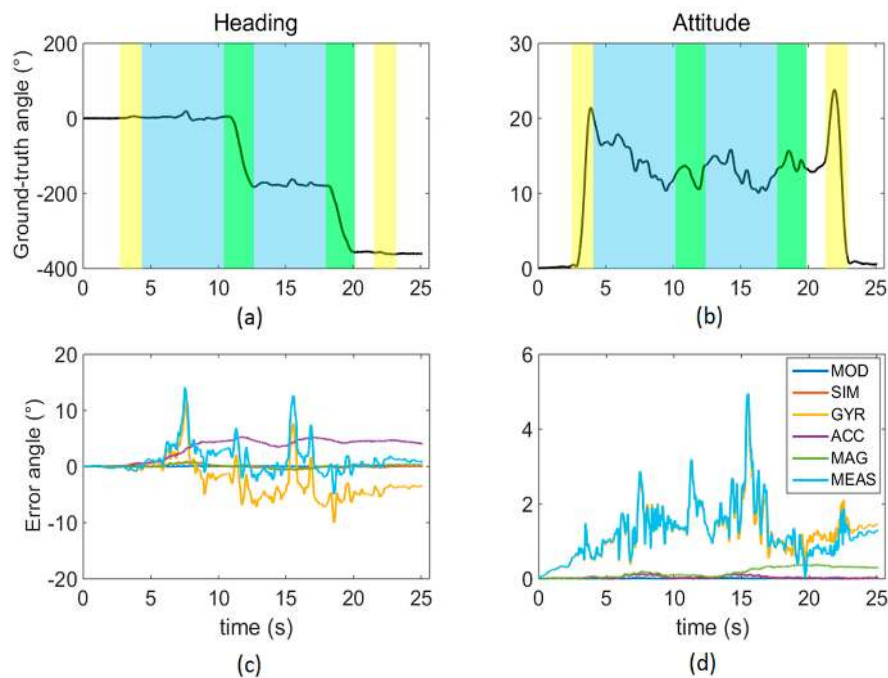


Figure 4. Heading (a) and attitude (b) ground-truth curves; heading (c) and attitude (d) error angles obtained for all the six scenarios considered. The colored bands in the upper row denote activities of sit-to-stand and stand-to-sit (**yellow**), walking (**blue**), and 180° turns around the cranio-caudal axis.

Figure 5 shows the RMS error values (medians and inter-quartile ranges) obtained for both Algorithms 1 and 2.

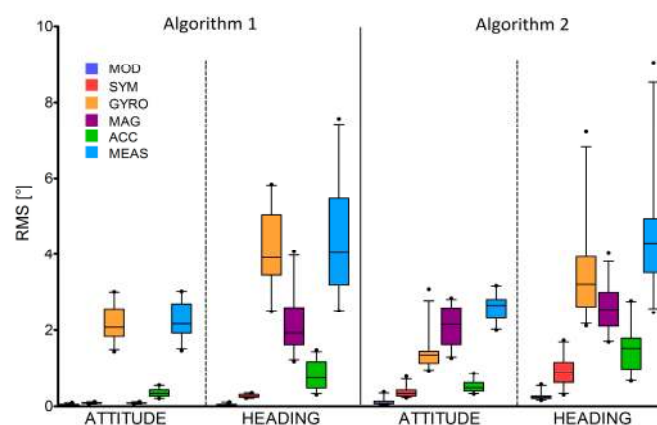


Figure 5. RMS_{head} and RMS_{att} median and inter quartile ranges obtained for the two considered SFAs and all the six tested scenarios.

For what concerns the intra-algorithm comparisons, the results of the one-way repeated measures ANOVA are reported in Table 2. A significant scenario effect was found for both heading and attitude errors and for both Algorithms 1 and 2. The scenario factor accounts for a minimum of 83% to a maximum of 97% of the overall orientation errors variance.

Table 2. Results of the one-way repeated measures ANOVA for both the attitude and heading errors and for each tested algorithm. Degrees of freedom for the effect (df_{scenario}) and for the error term (df_{error}) are reported together with F values, p values and partial eta squared (η^2).

Algorithm	MOP	df_{scenario}	df_{error}	F	p	η^2
1	RMS _{att}	1.13	25.93	193.95	<0.001	0.85
	RMS _{head}	3.09	71.21	661.30	<0.001	0.97
2	RMS _{att}	1.05	24.26	113.60	<0.001	0.83
	RMS _{head}	3.37	77.67	285.84	<0.001	0.92

The selection of the scenarios considered for the *post hoc* tests was done according to the questions presented in Section 2.6. In Table 3, the results of the statistical analysis are reported (mean differences between the errors obtained in the considered scenarios and significance).

Table 3. *Post-hoc* analysis: marginal differences between the scenarios indicated in the first and second column for both algorithms. Significant differences are indicated with an asterisk.

Tested Scenario	Testbed Scenario	Algorithm 1		Algorithm 2	
		Attitude	Heading	Attitude	Heading
GYR	SIM	2.11 *	3.82 *	1.03 *	2.61 *
	MEAS	−0.9 *	−0.19	−1.22 *	−0.98 *
ACC	SIM	0.25 *	0.54 *	0.15 *	0.61 *
	MEAS	−1.95 *	−3.47 *	−2.10 *	−2.99 *
MAG	SIM	0.00	1.93 *	1.76 *	1.66 *
	MEAS	−2.20 *	−2.08 *	−0.49 *	−1.94 *
MOD	SIM	−0.46 *	−0.22 *	−0.25 *	−0.64 *

The results of the Wilcoxon test are reported in Table 4. Significant differences were found between the two algorithms for both attitude and heading errors for the SIM scenario. On the other hand, for the MEAS scenario, the two algorithms provided significantly different results only for the attitude angle.

Table 4. Results (Z and p -values) of the comparison between Algorithms 1 and 2 for the SIM and MEAS scenarios. To improve the table readability/clarity, significant differences are also indicated with an asterisk.

Considered Scenario	Attitude		Heading	
	Z	p	Z	p
SIM	−4.286	<0.001 *	−4.286	<0.001 *
MEAS	−2.143	0.032 *	−0.829	0.407

4. Discussions

In the present study, a novel benchmarking method for SFAs performance assessment was presented and applied to a human movement analysis case study. In particular, the 3D pelvis orientation was estimated during a TUG test. A pool of hybrid scenarios, including both simulated and real MIMU data, was created and the combination of data associated to each scenario was fed to two different Kalman-based SFAs. The accuracy with which the pelvis orientation was estimated was then assessed by comparing the output of each algorithm with the reference output.

In Figure 4, the attitude curve exhibits two transitions at the beginning and the end of the test (about 3 s and 23 s), clearly corresponding to the sit-to-stand and stand-to-sit phases. The two 180°

turns are also visible in the heading plot. Interestingly, both the attitude and heading error trends do not seem to be related with any of the activities mentioned above.

For both SFAs, the orientation errors were, on average, 3° and 4° for the attitude and heading, respectively (Figure 5). These errors can be considered rather large given the short test time duration (less than 20 s). In addition, it is clear that, for the task being shown, the GYR errors are highly correlated with the MEAS ones, suggesting a crucial role of the gyroscope errors in determining the performance of the SFA considered. To reduce orientation estimation errors, SFAs may be revised and tuned and, to do so, the information about the effect of the different noise sources characterizing each MIMU sensor is fundamental. The proposed benchmarking method aimed at providing a contribution in this respect. The following considerations are a concrete example of the information that can be derived by applying the proposed method for the assessment of a generic SFA.

For what concerns the intra-algorithm evaluation, a significant scenario effect was found for both Algorithms 1 and 2. This result was expected as the scenarios consisted of different combinations of simulated and real sensor data. The cascaded *post-hoc* analysis between the SIM and GYR scenarios revealed that using the simulated gyroscope data instead of the measured ones produced a significant effect for both attitude and heading accuracy, regardless of the considered SFA (Table 3). This means that the errors characterizing the gyroscope have a detrimental effect on both SFAs performance. In addition, both SFAs were found to take advantage of the simulated accelerometer and magnetometer data to partially correct the gyroscope-related errors (GYR *vs.* MEAS). This was always true except for the heading estimation obtained using Algorithm 1 (Table 3). Such a result may be explained by focusing on the inner structure of this specific algorithm, for which the angular velocity obtained from the gyroscope drives the accelerometer and magnetometer data pre-processing. Algorithm 1 is, thus, particularly sensitive to gyroscope imperfections and inconsistencies. Conversely, for what concerns Algorithm 2, accelerometer and magnetometer data are used to explicitly correct the quaternion predicted according to the gyroscope measurement. Therefore, the update step performed in this algorithm appears to be effective in correcting the prediction errors due to the gyroscope when the accelerometer and magnetometer are ideal. In any case, the GYR scenario is clearly the closest to the MEAS scenario, which represents the upper bound for the SFA error. This means that, both SFAs are more sensitive (both in terms of attitude and heading) to the gyroscope issues than to the accelerometer and magnetometer ones.

Significant differences were reported also between the ACC scenario and both the SIM and MEAS scenarios. On the one hand, the imperfections affecting the real accelerometer measurements degrade significantly the orientation estimates with respect to the SIM scenario. On the other hand, it has to be noted that both SFAs successfully exploit the other simulated sensors to reduce the errors due to these imperfections, as highlighted by the comparison between the ACC and the MEAS scenarios. Interestingly, significant differences between ACC and SIM were found also for the heading estimation, indicating that, although gravity measurements (from the accelerometer) do not convey heading information, errors in the accelerometer data cause errors in the heading estimation. This is also supported by the results of the comparison between the MOD and SIM scenarios: for both SFAs and both heading and attitude, a significant difference was observed between the two scenarios, which only differ for the presence of external accelerations (accelerometer inconsistency [3]). These two results are in accordance with the existing literature [24], where heading notoriously represents the most difficult degree of freedom to be estimated with MIMU data [6]. In fact, it is well known that attitude errors imply an additional error on heading [24]. Unfortunately, no algorithm design can prevent this effect.

When considering the heading estimation using both Algorithms 1 and 2, the errors associated to the MAG scenario were significantly different from those obtained for both the SIM and MEAS scenarios. On the other hand, for Algorithm 1, when the attitude is considered, no significant difference was observed between the MAG and SIM scenarios. This result may be explained as follows: within Algorithm 1 the actual orientation estimation is performed by the TRIAD block (see Figure 3), which uses the estimated gravity as the first (more reliable) direction. As a result, magnetic inconsistencies

(due to ferromagnetic disturbances) are prevented from degrading significantly the attitude estimation. In other words, in Algorithm 1, the magnetometer readings are almost neglected for the attitude estimation. Conversely, for Algorithm 2, magnetic disturbances have a significant detrimental effect both on attitude and heading estimation, as magnetometer data are used for the attitude estimation as well. Given this result, it is recommendable to decouple the attitude estimation from the magnetometer measurements when designing SFA for orientation estimation targeted to indoor applications, where the hypothesis of uniform and constant magnetic field can be undermined.

With regard to the inter-algorithm comparison (Table 4), when considering the SIM scenario, a statistically different performance was exhibited by the two SFAs. In particular Algorithm 1 was characterized by smaller heading and attitude errors. This result can be explained by the different algorithm design and, in particular, by the fact that Algorithm 1 takes advantage of the linear approach to the Kalman filtering over the extended one implemented in Algorithm 2. On the other hand, this performance gap is much less evident when real sensor data are involved, as shown by the comparison of the two MEAS scenarios (Table 4). Even though a significant effect is reported for the attitude estimation in the MEAS scenarios, it should be more reasonably attributed to the exclusion of the magnetic measurements from the attitude estimation than to the linear estimation approach. In fact, no statistical differences were found for the heading estimation, for which the magnetometer is used by both methods. In other words, the different designs might indeed imply different estimation performances. However, these differences may be concealed by the inconsistencies and imperfections characterizing real data obtained from low grade MIMUs.

As a summary, the proposed benchmarking method has the benefit to allow for an improved understanding of the extent to which each sensor issue influences the final orientation estimate. This possibility is of paramount importance in the context of SFAs design to correctly consider the role of each sensor. As an example, by means of the proposed methodology the following considerations for the MIMU-based human motion tracking can be drawn: (1) the gyroscope errors appear to be the main error source for both the SFAs considered; (2) the processing of accelerometer data proposed in Algorithm 1 is promising because it reduces the detrimental effect of the external acceleration; (3) using the magnetometer data for the heading estimation only leads to more accurate attitude estimates. These considerations can then be used to improve the existing designs/tuning settings in accordance to the needs of the specific application.

5. Conclusions

In conclusion, the methodology presented in this paper allows to gain insight into the working principles of a generic sensor fusion algorithm. The proposed data hybridization process consists in combining real and ideal sensor data with the aim of addressing the main strengths and weaknesses of an SFA. Moreover, the results obtained with this methodology when applied to multiple SFAs allow highlighting their specific behavior with respect to different input data issues. The proposed benchmarking method was applied to a typical sensor fusion context in human movement analysis, namely 3D orientation estimation through MIMU data. The potential of the proposed methodology was thus exploited to reveal the main issues involved in this specific application context.

Acknowledgments: This work has been supported by funds from the Italian Ministry of Education and Research (PRIN 2010R277FT).

Author Contributions: Gabriele Ligorio contributed to: original concept; study conception and design; data processing; data analysis and interpretation; manuscript drafting and revision. Elena Bergamini contributed to: study conception and design; data acquisition; statistical analysis; data analysis and interpretation; manuscript drafting and revision. Ilaria Pasciuto and Giuseppe Vannozzi contributed to: manuscript drafting and critical revision. Aurelio Cappozzo contributed to: supervision of the work; manuscript critical revision. Angelo Maria Sabatini contributed to: supervision of the work; data analysis and interpretation; manuscript drafting and revision; manuscript critical revision.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In this appendix the two orientation estimators (Algorithms A1 and A2) evaluated and compared in this paper are detailed with a CLRS description, where:

- \mathbf{x}_k and $\hat{\mathbf{x}}_k$ are, respectively, the real and estimated value of the state vector \mathbf{x} ;
- $\hat{\mathbf{x}}_k^a$ and $\hat{\mathbf{x}}_k^b$ are, respectively, the *a priori* and *a posteriori* estimate of the state vector \mathbf{x} at the k -th time instant;
- \mathbf{h}^n and \mathbf{h}^b are the representation of the Earth's magnetic field in the $\{\mathbf{n}\}$ and $\{\mathbf{b}\}$ reference frames, respectively;
- \mathbf{g}^n and \mathbf{g}^b are the representation of the Earth's magnetic field in the $\{\mathbf{n}\}$ and $\{\mathbf{b}\}$ reference frames, respectively;
- $^s\sigma$, $^a\sigma$ and $^m\sigma$ are the standard deviations of the gyroscope, accelerometer and the magnetometer measurement, respectively;
- $^b\sigma$ is a tuning parameter encoding the expected severity of the magnetic disturbances in method B;
- c_a and c_b are the Gauss-Markov parameters of the prediction models in method A;
- \mathbf{H} , \mathbf{Q} and \mathbf{R} are, respectively, the extraction matrix, the process noise matrix and the measurement noise matrix in the Kalman filter;
- T_s is the sampling time;
- \times represents the cross product operator;
- \otimes represents the quaternion product operator;
- \mathbf{q}^{bn} and \mathbf{R}^{bn} represents, respectively, the quaternion and the rotation matrix encoding the orientation from the $\{\mathbf{n}\}$ reference frame to the $\{\mathbf{b}\}$ reference frame.
- \mathbf{I}_n and $\mathbf{0}_n$ are the $n \times n$ identity and null matrix;

- $[\mathbf{p} \times]$ is the cross product matrix for the vector $\mathbf{p} = [p_1 \ p_2 \ p_3]^T$,

$$[\mathbf{p} \times] = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix};$$

- $\Omega(\mathbf{p}) = \begin{bmatrix} -[\mathbf{p} \times] & \mathbf{p} \\ -\mathbf{p}^T & 0 \end{bmatrix};$

- $\Xi(\mathbf{q}) = \begin{bmatrix} [{}^v\mathbf{q} \times] + q_4 \cdot \mathbf{I}_3 \\ {}^v\mathbf{q}^T \end{bmatrix}$ with $\mathbf{q} = [q_1 \ q_2 \ q_3 \ q_4]^T = [{}^v\mathbf{q} \ q_4]^T$ being a generic unit quaternion;

- $\Psi(\mathbf{q}, \mathbf{p}) = \frac{\partial \left(\mathbf{q} \otimes \begin{bmatrix} \mathbf{p} \\ 0 \end{bmatrix} \otimes \mathbf{q}^{-1} \right)}{\partial (\mathbf{q})}$

Algorithm A1

for each measured sample from the sensors **do**

$\omega_k^b \leftarrow$ gyroscope sample

Prediction

Predict gravity and acceleration (KF_{acc})

$$\bar{acc}\mathbf{x}_k = \begin{bmatrix} -\hat{\mathbf{g}}_k^b \\ -\hat{\mathbf{a}}_k^b \end{bmatrix} = \begin{bmatrix} \exp(-[\omega_{k-1}^b \times] \cdot T_s) & \mathbf{0}_3 \\ \mathbf{0}_3 & c_a \cdot \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} +\hat{\mathbf{g}}_{k-1}^b \\ +\hat{\mathbf{a}}_{k-1}^b \end{bmatrix}$$

Process noise covariance matrix (KF_{acc})

$$acc\mathbf{Q}_{k-1} = \begin{bmatrix} [-\hat{\mathbf{g}}_k^b \times] & \mathbf{0}_3 \\ \mathbf{0}_3 & c_b \cdot \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} \mathbf{I}_3 \cdot (\sigma T_s)^2 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} [-\hat{\mathbf{g}}_k^b \times] & \mathbf{0}_3 \\ \mathbf{0}_3 & c_b \cdot \mathbf{I}_3 \end{bmatrix}^T$$

Predict Earth's magnetic field and disturbances (KF_{mag})

$$\bar{mag}\mathbf{x}_k = \begin{bmatrix} -\hat{\mathbf{h}}_k^b \\ -\hat{\mathbf{d}}_k^b \end{bmatrix} = \begin{bmatrix} \exp(-[\omega_{k-1}^b \times] \cdot T_s) & \mathbf{0}_3 \\ \mathbf{0}_3 & c_a \cdot \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} +\hat{\mathbf{h}}_{k-1}^b \\ +\hat{\mathbf{d}}_{k-1}^b \end{bmatrix}$$

Process noise covariance matrix (KF_{mag})

$$mag\mathbf{Q}_{k-1} = \begin{bmatrix} [-\hat{\mathbf{h}}_k^b \times] & \mathbf{0}_3 \\ \mathbf{0}_3 & c_b \cdot \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} \mathbf{I}_3 \cdot (\sigma T_s)^2 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} [-\hat{\mathbf{h}}_k^b \times] & \mathbf{0}_3 \\ \mathbf{0}_3 & c_b \cdot \mathbf{I}_3 \end{bmatrix}^T \text{ Update:}$$

$$\mathbf{H} =_{mag} \mathbf{H} =_{acc} \mathbf{H} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{I}_3 \end{bmatrix}$$

Measurement update (KF_{acc})

$acc\mathbf{y}_k \leftarrow$ accelerometer sample

$$acc\mathbf{R} = \mathbf{I}_3 \cdot \sigma^2$$

$$+_{acc}\hat{\mathbf{x}}_k \leftarrow \text{KalmanUpdate}(\bar{acc}\mathbf{x}_k, acc\mathbf{Q}_{k-1}, acc\mathbf{y}_k, \mathbf{H}, acc\mathbf{R})$$

Measurement update (KF_{mag})

$mag\mathbf{y}_k \leftarrow$ magnetometer sample

$$mag\mathbf{R} = \mathbf{I}_3 \cdot \sigma^2$$

$$+_{mag}\hat{\mathbf{x}}_k \leftarrow \text{KalmanUpdate}(\bar{acc}\mathbf{x}_k, mag\mathbf{Q}_{k-1}, mag\mathbf{y}_k, \mathbf{H}, mag\mathbf{R})$$

Orientation computation (TRIAD algorithm)

Normalize $+\hat{\mathbf{g}}_k^b$ and $+\hat{\mathbf{h}}_k^b$ to one

$$\mathbf{M}_{obs} = \begin{bmatrix} +\hat{\mathbf{g}}_k^b & +\hat{\mathbf{g}}_k^b \times +\hat{\mathbf{h}}_k^b & +\hat{\mathbf{g}}_k^b \times (+\hat{\mathbf{g}}_k^b \times +\hat{\mathbf{h}}_k^b) \end{bmatrix}$$

$$\mathbf{M}_{ref} = \begin{bmatrix} \mathbf{g}^n & \mathbf{g}^n \times \mathbf{h}^n & \mathbf{g}^n \times (\mathbf{g}^n \times \mathbf{h}^n) \end{bmatrix}$$

$$\hat{\mathbf{q}}_k^{bn} = \text{matrixToQuaternion}(\mathbf{M}_{ref}\mathbf{M}_{obs}^T)$$

end for

Algorithm A2

for each measured sample from the sensors **do**

$\omega_k^b \leftarrow$ gyroscope sample

Prediction

Predict the quaternion and the magnetic disturbance

$$-\hat{\mathbf{x}}_k = \begin{bmatrix} -\hat{\mathbf{q}}_k^{bn} \\ -\hat{\mathbf{b}}_k^n \end{bmatrix} = \begin{bmatrix} \exp\left(\Omega\left(\omega_{k-1}^b\right)T_s\right) & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} +\hat{\mathbf{q}}_{k-1}^{bn} \\ +\hat{\mathbf{b}}_{k-1}^n \end{bmatrix}$$

Process noise covariance matrix

$$\mathbf{Q}_{k-1} = \begin{bmatrix} \left(\frac{g\sigma T_s}{2}\right)^2 \Xi\left(+\hat{\mathbf{q}}_{k-1}^{bn}\right)\mathbf{I}_3\Xi\left(+\hat{\mathbf{q}}_{k-1}^{bn}\right)^T & \mathbf{0}_{4\times 3} \\ \mathbf{0}_{3\times 4} & b\sigma^2 \cdot \mathbf{I}_3 \end{bmatrix}$$

Update:

compute $-\hat{\mathbf{R}}^{bn} = \text{quaternionToMatrix}(-\hat{\mathbf{q}}^{bn})$

$$\mathbf{y}_k = \begin{bmatrix} \text{mag}\mathbf{y}_k \\ \text{acc}\mathbf{y}_k \end{bmatrix}, \quad \begin{array}{l} \text{mag}\mathbf{y}_k \leftarrow \text{magnetometer sample} \\ \text{acc}\mathbf{y}_k \leftarrow \text{accelerometer sample} \end{array}$$

Sensor data validation

$$\text{acc}\mathbf{R} = \begin{cases} \mathbf{I}_3 \cdot a \sigma^2, & \text{if } \left| \text{acc}\mathbf{y} - -\hat{\mathbf{R}}^{bn} \mathbf{g}^n \right| <_{\text{acc}} \varepsilon \\ \infty, & \text{otherwise} \end{cases}$$

$$\text{mag}\mathbf{R} = \begin{cases} \mathbf{I}_3 \cdot m \sigma^2, & \text{if } \left| \text{mag}\mathbf{y} - -\hat{\mathbf{R}}^{bn} \left(\mathbf{h}^n - -\hat{\mathbf{b}}_k^n \right) \right| <_{\text{mag}} \varepsilon \\ \infty, & \text{otherwise} \end{cases}$$

$$\mathbf{R} = \begin{bmatrix} \text{mag}\mathbf{R} \\ \text{acc}\mathbf{R} \end{bmatrix}$$

Jacobian calculation

$$\mathbf{H} = \begin{bmatrix} \Psi\left(-\hat{\mathbf{q}}_k^{bn}, \mathbf{h}^n + -\hat{\mathbf{b}}_k^n\right) & \mathbf{R}^{bn} \\ \Psi\left(-\hat{\mathbf{q}}_k^{bn}, \mathbf{g}^n\right) & \mathbf{0}_3 \end{bmatrix}$$

$$+\hat{\mathbf{x}} \leftarrow \text{KalmanUpdate}\left(-\hat{\mathbf{x}}, \mathbf{Q}_{k-1}, \mathbf{y}_k, \mathbf{H}, \mathbf{R}\right)$$

end for

References

1. Corke, P.; Lobo, J.; Dias, J. An Introduction to Inertial and Visual Sensing. *Int. J. Robot. Res.* **2007**, *26*, 519–535. [CrossRef]
2. Novak, D.; Riener, R. A survey of sensor fusion methods in wearable robotics. *Robot. Auton. Syst.* **2015**, *73*, 155–170. [CrossRef]
3. Khaleghi, B.; Khamis, A.; Karray, F.O.; Razavi, S.N. Multisensor data fusion: A review of the state-of-the-art. *Inf. Fusion* **2013**, *14*, 28–44. [CrossRef]
4. Weiss, S.; Scaramuzza, D.; Siegwart, R. Monocular-SLAM-based navigation for autonomous micro helicopters in GPS-denied environments. *J. Field Robot.* **2011**, *28*, 854–874. [CrossRef]
5. Lu, Y.; Michaels, J.E. Feature extraction and sensor fusion for ultrasonic structural health monitoring under changing environmental conditions. *IEEE Sens. J.* **2009**, *9*, 1462–1471.

6. Bergamini, E.; Ligorio, G.; Summa, A.; Vannozzi, G.; Cappozzo, A.; Sabatini, A.M. Estimating Orientation Using Magnetic and Inertial Sensors and Different Sensor Fusion Approaches: Accuracy Assessment in Manual and Locomotion Tasks. *Sensors* **2014**, *14*, 18625–18649. [[CrossRef](#)] [[PubMed](#)]
7. Mazomenos, E.; Biswas, D.; Cranny, A.; Rajan, A.; Maharatna, K.; Achner, J.; Klemke, J.; Jobges, M.; Ortmann, S.; Langendoerfer, P. Detecting Elementary Arm Movements by Tracking Upper Limb Joint Angles with MARG Sensors. *IEEE J. Biomed. Health Inform.* **2015**. [[CrossRef](#)] [[PubMed](#)]
8. Sabatini, A.M. Inertial sensing in biomechanics: A survey of computational techniques bridging motion analysis and personal navigation. In *Computational Intelligence for Movement Sciences: Neural Networks and Other Emerging Techniques*; Idea Group Publishing: Hershey, PA, USA, 2006; pp. 70–100.
9. Sabatini, A.M. Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors* **2011**, *11*, 1489–1525. [[CrossRef](#)] [[PubMed](#)]
10. Ligorio, G.; Sabatini, A.M. A Linear Kalman Filtering-Based Approach for 3D Orientation Estimation from Magnetic/Inertial Sensors. In Proceedings of the IEEE International Conference on Multisensor Fusion and Information Integration, San Diego, CA, USA, 14–16 September 2015.
11. Harada, T.; Mori, T.; Sato, T. Development of a tiny orientation estimation device to operate under motion and magnetic disturbance. *Int. J. Robot. Res.* **2007**, *26*, 547–559. [[CrossRef](#)]
12. Choukroun, D.; Bar-Itzhack, I.Y.; Oshman, Y. Novel quaternion Kalman filter. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 174–190. [[CrossRef](#)]
13. Madgwick, S.O.; Harrison, A.J.; Vaidyanathan, R. Estimation of IMU and MARG orientation using a gradient descent algorithm. In Proceedings of the 2011 IEEE International Conference on Rehabilitation Robotics (ICORR), Zurich, Switzerland, 29 June–1 July 2011; pp. 1–7.
14. Mahony, R.; Hamel, T.; Pflimlin, J.-M. Nonlinear Complementary Filters on the Special Orthogonal Group. *IEEE Trans. Autom. Control* **2008**, *53*, 1203–1218. [[CrossRef](#)]
15. Titterton, D.; Weston, J.L. *Strapdown Inertial Navigation Technology*; IET: Stevenage, UK, 2004.
16. Bancroft, J.B.; Lachapelle, G. Estimating MEMS gyroscope g-sensitivity errors in foot mounted navigation. In Proceedings of the Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS), Helsinki, Finland, 3–4 October 2012; pp. 1–6.
17. Strauss, O. Filtering and fusing compass and gyrometer data using Guess filter. In Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, Barcelona, Spain, 1–5 July 1997.
18. Winter, D.A. *Biomechanics and Motor Control of Human Movement*; John Wiley & Sons: New York, NY, USA, 2009.
19. Chardonnes, J.; Favre, J.; Aminian, K. An effortless procedure to align the local frame of an inertial measurement unit to the local frame of another motion capture system. *J. Biomech.* **2012**, *45*, 2297–2300. [[CrossRef](#)] [[PubMed](#)]
20. Ligorio, G.; Sabatini, A.M. A Simulation Environment for Benchmarking Sensor Fusion-Based Pose Estimators. *Sensors* **2015**, *15*, 32031–32044. [[CrossRef](#)] [[PubMed](#)]
21. Shuster, M.D. A survey of attitude representations. *Navigation* **1993**, *8*, 439–517.
22. Black, H.D. A passive system for determining the attitude of a satellite. *AIAA J.* **1964**, *2*, 1350–1351. [[CrossRef](#)]
23. Faber, G.S.; Chang, C.-C.; Rizun, P.; Dennerlein, J.T. A novel method for assessing the 3-D orientation accuracy of inertial/magnetic sensors. *J. Biomech.* **2013**, *46*, 2745–2751. [[CrossRef](#)] [[PubMed](#)]
24. Caruso, M.J. Applications of magnetic sensors for low cost compass systems. In Proceedings of the Position Location and Navigation Symposium, San Diego, CA, USA, 13–16 March 2000; pp. 177–184.

