Article

A Multi-Modal Under-Sensorized Wearable System for Optimal Kinematic and Muscular Tracking of Human Upper Limb Motion

Paolo Bonifati [1,*,†] , Marco Baracca [1,†] , Mariangela Menolotto [1], Giuseppe Averta [2] and Matteo Bianchi [1]

1   Research Center "E. Piaggio", Department of Information Engineering, University of Pisa,
    Largo Lucio Lazzarino 1, 56126 Pisa, Italy
2   Department of Control and Computer Engineering, Politecnico di Torino, 10129 Torino, Italy
*   Correspondence: paolo.bonifati@phd.unipi.it
†   These authors contributed equally to this work.

Abstract: Wearable sensing solutions have emerged as a promising paradigm for monitoring human musculoskeletal state in an unobtrusive way. To increase the deployability of these systems, considerations related to cost reduction and enhanced form factor and wearability tend to discourage the number of sensors in use. In our previous work, we provided a theoretical solution to the problem of jointly reconstructing the entire muscular-kinematic state of the upper limb, when only a limited amount of optimally retrieved sensory data are available. However, the effective implementation of these methods in a physical, under-sensorized wearable has never been attempted before. In this work, we propose to bridge this gap by presenting an under-sensorized system based on inertial measurement units (IMUs) and surface electromyography (sEMG) electrodes for the reconstruction of the upper limb musculoskeletal state, focusing on the minimization of the sensors' number. We found that, relying on two IMUs only and eight sEMG sensors, we can conjointly reconstruct all 17 degrees of freedom (five joints, twelve muscles) of the upper limb musculoskeletal state, yielding a median normalized RMS error of 8.5% on the non-measured joints and 2.5% on the non-measured muscles.

Keywords: human multimodal motion tracking; optimal design; Sensor Fusion; IMUs; sEMG sensors; upper limb; wearable sensing

1. Introduction

The evaluation of the musculoskeletal state of the human body is crucial for different applications, such as rehabilitation and assistive technologies [1], sportsmen monitoring [2,3] and human-robot interaction and collaboration [4]. Such a monitoring is also important to prevent possible work-related musculoskeletal disorders, providing tools for a proper ergonomics evaluation [5–7] informed by suitably devised bio-mechanical models [8].

Considering the degrees of freedom (DoFs) of the human body, i.e., joints and muscular sites, a correct tracking of human kinematics and muscular activity would require the acquisition of a large amount of data and the usage of many sensors [9]. To record muscle activation, the standard solution is surface electromyography (sEMG), which relies on the usage of electrodes fastened on the skin that measure the electric signal (expressed in mV) produced by muscles. For kinematic measures, instead, the gold standard has been traditionally provided by optical systems, which can monitor human body motion by recording the 3D position in the time of active or passive optical markers. These systems have been proved to be efficient and reliable, but they come with limitations of the operating space. Furthermore, occlusions can also occur, thus affecting the overall reconstruction performance. This problem also affects other marker-less, camera-based methods that have been proposed [10]. A solution to address the problem of environmental occlusion

was presented in [11], where the authors exploited radio signals to estimate human pose through walls. However, this approach cannot be generalized to any distance from the sensor, or any type of occlusion, e.g., induced by the presence of other people.

Wearable solutions have emerged as a promising paradigm to enable ecological monitoring, overcoming the workspace limits that affect camera-based methods. Ergonomics and form-factor related considerations tend to discourage the usage of cumbersome sensors. Under this regard, inertial measurement unit (IMU)-based approaches have found fertile ground for kinematic tracking, thanks to their compact design and reduced costs [12,13].

However, to obtain a full biomechanical assessment of the human body, kinematic information is not sufficient but it should be complemented with the recording of muscular activation, e.g., to correctly evaluate the fatigue level of the user during task execution [14–16]. Simultaneous acquisition and fusion of muscular and kinematic information have been proposed, e.g., in [17], where measurements from IMUs and mechanomyography were exploited for classifying different actions of the lower limb and for evaluating pathological state. Of note, wearable solutions (eventually complemented, in some cases, by cost considerations) tend to discourage the usage of many sensors mounted on the body, which could negatively impact the form factor and the wearability of the device [18]. A possible approach to tackle this issue is to exploit the covariation schemes between functional elements or DoFs of our body, usually named as *motor synergies* [19]. Indeed, several works demonstrated the existence of correlation patterns between different joints and/or muscles in the upper [20–23] and lower limb [24,25]. The underlying idea is that the actuation of a large number of DoFs can be described as a linear combination of a smaller number of generators. In terms of actuation schemes, this concept has been profitably exploited in robotics for the design [26], planning [27,28] and control [29] of anthropomorphic devices, with a special focus on robotic hands and manipulators. In all these cases, a small number of independent actuation variables can be combined to drive a larger number of DoFs in a human-like fashion.

Interestingly, the same paradigm can also be used to inform simplified sensing strategies of human motion. In [30], we demonstrated that it is possible to complement scarce and noisy sensory information on hand grasping posture by fusing it with a priori data through minimum variance estimation (MVE). A priori data represented the most frequent human grasping postures organized in terms of interjoint covariation patterns. In [31], we further built on this approach and identified which were the optimal hand joints that yield the minimization in average of the reconstruction error, exploiting the minimization of the a posteriori covariance matrix. These results allowed us to design a wearable sensing glove to reconstruct the hand pose, relying on a lower number of sensors [32]. However, these approaches are based on the assumption that the a priori information is related to static postures, and their application to the estimation of temporal trajectories cannot be performed in a straightforward manner. Additionally, it is hard to develop a trustworthy estimation of the covariance matrix from heterogeneous data due to the concurrent reconstruction of multimodal motion-related data (such as joint angles and EMG signals) [33]. In [34], we proposed to generalize these methods for the estimation of multi-modal time-varying data of the upper limb. The method built upon the existence of covariation patterns in human upper limb motions, as we demonstrated in [23] and the usage of functional analysis for reconstructing the whole trajectory over time and estimating the covariance matrix. In brief, a base of functional Principal Components (fPCs), derived in advance from a collection of upper limb joint motion profiles of daily living activities, was employed to map the temporal measurements of a reduced number of joints and muscles on the extended state space of weights and average trajectories/muscles envelopes. The state missing part was then reconstructed using MVE. The temporal evolution of the entire muscle-skeletal system is then appropriately integrated with the estimated extended state.

However, in [34], the analysis was performed assuming as state variables the joint angular values and the muscle envelopes, while the non-linear mapping between sen-

sors and state variables was not considered. In this paper, we build upon our previous work and extend the method to design an under-sensorized wearable system for multi-modal acquisition of human upper limb trajectories. We assume to have at disposal IMUs for kinematic recording and surface sEMGs for muscular activity acquisition, and that their number is not in a bijective relation with all the DoFs used to describe the whole muscle-skeletal status. We generalize the optimal sensing setup identified in [34] to the more challenging case in which one sensor may record the activity of multiple DoFs. Indeed, since the goal is now to reduce the number of employed sensor elements, instead of selecting the single optimal degrees of freedom, i.e., the ones that are associated with a reduced estimation uncertainty, our targeted optimal joint angles are those that enable a compromise between optimal reconstruction and the minimization of the sensing resource in use. To target both objectives, we select as measures the shoulder joints. In this way, we minimize the differences with respect to the optimal setup reported in [34]. Finally, we build a real prototype of an optimal under-sensorized setup for upper limbs (i.e., which has a number of elements lower than the number required to measure all states of the system), with only two IMUs to retrieve angles from the shoulder by implementing an Unscented Kalman Filter (UKF). We integrated these measurements with the optimal sEMGs identified in [34], discarding the other ones, and using a commercially available fully sensorized solution (i.e., Xsens) to have a ground truth for result comparison. Extensive tests on a dataset collected with our framework demonstrate that our method can effectively compensate for missing recordings (corresponding to two out of five joint angles and four out of twelve sEMG signals), with minimum impact on the estimation error, achieving a median normalized RMS error of 8.5% on the non-measured joints and of 2.5% on the non-measured EMGs.

The paper is organized as follows: we first summarize the theory underpinning our optimization method and its application to our case, with the UKF implementation for retrieving shoulder angles; then, we discuss the experimental setup for data acquisition and system testing, and the results.

2. Methods

2.1. Theoretical Foundations: Minimum Variance Estimation (MVE)

Here we briefly summarize the results in [34]. The idea is to translate the recorded movements into a static representation, use it to obtain the *a priori* covariance matrix, perform the estimation and then re-express the movements in the temporal domain. To do this, we define three separate phases in this method: encoding, estimation and decoding. The procedure is briefly depicted in Figure 1.
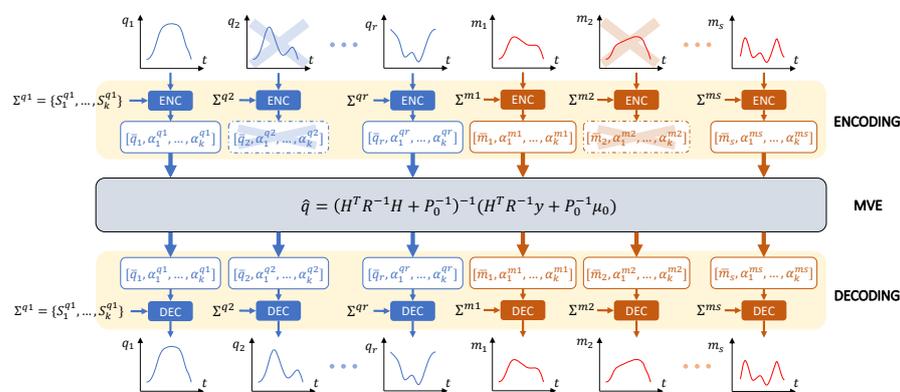

Figure 1. Schematic flow of the estimation procedure. First temporal signals are mapped on the weight vector through the fPC bases (*Encoding*). After that, *Minimum Variance Estimation* (MVE) fuses the encoded measures with *a priori* knowledge to estimate for the missing part of measures. In the end, the estimated weight vector is converted back to the temporal domain (*Decoding*).

2.1.1. Encoding and Decoding Phases: Functional Principal Component Analysis

Functional Principal Component Analysis (fPCA) is a statistical method to identify functional primitives from time-varying data. In this section, we will provide a brief introduction to the theory, while werefer to [35] for more details. For the sake of simplicity, since each DoF can be analyzed separately from the others with this method, the equations will be defined for a single joint. Let us consider N independent observations of joint temporal evolution $q_1(t), \ldots, q_N(t)$ with $t \in [0, 1]$. A generic motion can be decomposed as a weighted sum of basis elements $S_i(t)$, known as *functional Principal Components* (fPCs):

$$q(t) \simeq \bar{q} + S_0(t) + \sum_{i=1}^{s_{max}} \alpha_i S_i(t) \tag{1}$$

where $\bar{q}$ is the average value of the joint, $S_0(t)$ is the average trajectory across all the trajectories in the dataset, α_i is the weight associated with the ith basis element $S_i(t)$ and s_{max} is the number of basis elements. The output of fPCA is a basis of functions $\{S_1(t), \ldots, S_{s_{max}}(t)\}$ which maximizes the explained variances of joint motions throughout the whole dataset. For more detail on how these fPCs can be extracted, we refer the interested reader to [35].

This decomposition can be done for each DoF of the considered system, regardless of whether it is a kinematic or muscular measure, and it allows us to translate the trajectories from the time domain to the fPCs weight domain. Then, it is possible to represent movments that an extended state x_e , which does not depend on time, to represent movements. Given M degrees of freedom and using k fPCs for the decomposition, the extended state, from which we can compute the covariance matrix P_0, can be defined as:

$$x_e = \begin{bmatrix} \bar{x}_1 \; \alpha_1^{x_1} \; \ldots \; \alpha_k^{x_1} \; | \; \ldots \; | \; \bar{x}_M \; \alpha_1^{x_M} \; \ldots \; \alpha_k^{x_M} \end{bmatrix}^T \tag{2}$$

where x_i is the generic *i-th* degree of freedom. This new state definition is the output of the encoding phase and it will be used as the state of the MVE.

When performing fPCA to decompose a signal, the noise is usually represented by the higher-order components. Indeed, the fPC decomposition allows truncating this basis to include only a few elements ordered based on the variance they can account for, giving an additional tool to minimize the effect of noise in the a priori covariance matrix, which will be introduced in the next section. In our work, we used the first 7 functional Principal Components out of 10, which can account for a cumulative variance greater than 95% for each DoF.

Regarding the decoding phase, given the estimation of the extended state $\hat{x}_e$ provided by the MVE, we can return to the temporal domain by combining the fPCs through (1).

2.1.2. Estimation Phase: Minimum Variance Estimation

The Minimum Variance Estimation (MVE) approach is an algorithm that leverages on the information of a set of *a priori* observations, organized in terms of mean μ_0 and covariance matrix P_0, to estimate missing or noisy measurements. In the following, we will briefly describe this method, while we address to [30] for more details.

Considering a vector of measures $y \in \mathbb{R}^d$ provided by a selection of d sensors, and assuming a linear relationship between the state variables $x \in \mathbb{R}^l$ and the measures y, then $y = Hx + \nu$, where $H \in \mathbb{R}^{d,l}$ is a full row rank measurement matrix and ν is the measurement noise. The goal is to estimate x given y when $d < l$. If the number of realizations of x (collected in a matrix of *a priori* $X \in \mathbb{R}^{l,N}$) is large enough, the covariance matrix results:

$$P_0 = \frac{(X - \bar{x})(X - \bar{x})^T}{N - 1} \tag{3}$$

where $\bar{x}$ is a matrix whose columns contain the average μ_0 of X. Given P_0, the best estimate $\hat{x}$ of x is the vector that solves the following optimization problem:

$$\hat{x} = \text{argmin} \, \frac{1}{2}(x - \mu_0)^T P_0^{-1}(x - \mu_0). \tag{4}$$

Assuming that ν is the zero mean Gaussian noise with covariance matrix R, the solution of (4) can be found in a closed form as:

$$\hat{x} = (H^T R^{-1} H + P_0^{-1})^{-1}(H^T R^{-1} y + P_0^{-1}\mu_0). \tag{5}$$

We can also define the *a posteriori* covariance matrix, which contains the information regarding the uncertainty of the associated state estimation, as:

$$P_P = (H^T R^{-1} H + P_0^{-1})^{-1} \tag{6}$$

Its maximum eigenvalue is a measure of the estimation uncertainty and its dependence on the selection matrix H allows us to link the quality of the estimation with the sensor placement. Hence, we can set up the following optimization problem to search for the best selection matrix H_{opt} given a certain number of sensors:

$$H_{opt} = \underset{H}{\text{argmin}} \, \sigma_{max}(P_P(H)) \tag{7}$$

There are different ways to solve this optimization. However, in our case, we have to preserve the particular structure of the selection matrix. Indeed, the matrix H is composed by squared blocks H_i of dimension $k + 1$, each of which is a diagonal matrix corresponding to the average signal and the first k fPC coefficients of the *i-th* degree of freedom, which represent the extended state in (2). To deal with this constraint, in our previous work [34], we used a genetic algorithm.

2.2. Musculoskeletal Model and Sensor Choice

We considered the same arm muscles (shown in Figure 2) and the same kinematic model (represented in Figure 3) composed of three rotational joint for the shoulder and two for the elbow reported in [34].
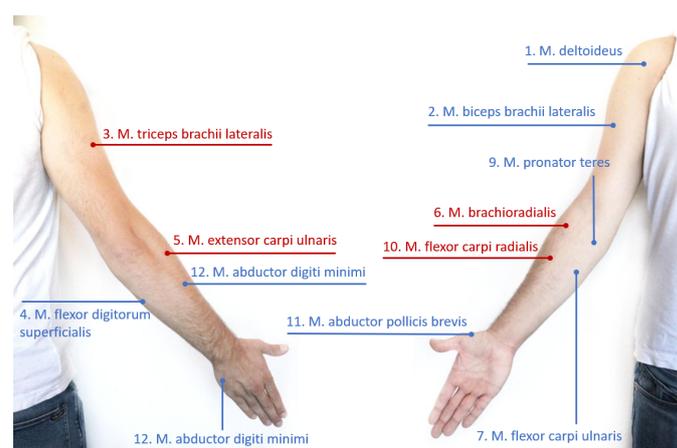

Figure 2. EMG sensor placement in accordance with SENIAM recommendations (back and front views of the right arm). In blue, the muscles used as measures in the MVE algorithm; in red, the estimated muscles.
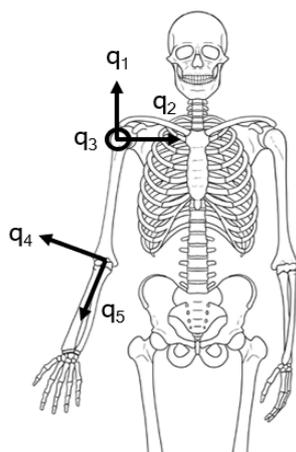
Figure 3. Kinematic model of the human arm (the angle q_3 is directed outwards).

In [34] the authors demonstrated that a good estimation of the biomechanical state of the arm can be reached measuring 3 joint angles (q_1, q_3, q_4 in Figure 3) and 8 muscular activation signals (indices 1, 2, 4, 7, 8, 9, 11, 12 in Figure 2). While the muscles optimal selection can be easily translated in the optimal sEMG sensor placement, for the kinematic measurements this is not necessarily true, since IMUs can capture the motion of several DoFs, depending on their placement. Indeed, usually two IMUs are placed before and after the anatomical articulation to estimate the joint angles of the kinematic model. To implement the results obtained in [34], a minimum number of 3 IMUs (one on the shoulder, one on the arm and one on the forearm) would be required. Since we are not assuming to measure every single joint independently from each other, moving from a discrete optimization to a continuous one, our goal is now to reduce the number of sensor elements while maximizing the lowest eigenvalue of the *a posteriori* covariance matrix P_p. Therefore, the idea is to select a sub-optimal set of joint angles (i.e., the ones of the shoulder q_1, q_2, q_3), which differs from the optimal case for just one DoF, but it requires only two IMUs for sensing.

2.3. Unscented Kalman Filter for Joint Angles Estimation via IMUs

Since the kinematic state of the upper-limb, and in particular the joint angles q and joint angular velocities $\dot{q}$, cannot be directly measured, a possible solution is based on an Unscented Kalman Filter (UKF) [36], which fuses the information given by a kinematic model of the arm with the measures of gyroscopes and accelerometers collected by two IMUs. Furthermore, the integration of magnetic field measures allows us to avoid the drifting behavior of the inertial sensors, which drastically limits the performance of the estimator.

Since we are solely interested in the measurement of the shoulder angles, from now on we can define the shoulder joint vector as $q = \begin{bmatrix} q_1, & q_2, & q_3 \end{bmatrix}^T$. The state space model of our UKF is based on the state $x(k) = \begin{bmatrix} q(k), & \dot{q}(k) \end{bmatrix}^T$, which contains the shoulder joints angles and the respective joint angular velocities at time k. The dynamic model of the i-th joint angle can be described with a first-order approximation as:

$$\begin{cases} q_i(k+1) = q_i(k) + \dot{q}_i(k) \cdot \Delta T + w_q(k) \\ \dot{q}_i(k+1) = \dot{q}_i(k) + w_{\dot{q}}(k) \end{cases} \tag{8}$$

where ΔT is the sampling time and the state is modelled as a random walk with Gaussian white noises w_q and $w_{\dot{q}}$.

The definition of the measurement model is based on the relationship between the inertial and magnetic field variables ω, a and m in the frames attached to the scapula IMU $\{S_R\}$ and the arm IMU $\{A_R\}$, passing through each pair of consecutive Denavit-Hartenberg frames $\{i\}$ and $\{i+1\}$. Assuming that the only value measured by the accelerometers is

the gravitational acceleration (i.e., the linear acceleration of the IMU and the Coriolis and centripetal accelerations are negligible) and that the two magnetometers are affected by the same disturbances, it is possible to write:

$$
\begin{cases}
\omega_{i+1}^{i+1} = R_{i+1,i} \left(\omega_i^i + z_i \cdot \dot{\theta}_{i+1} \right) \\
a_{i+1}^{i+1} = R_{i+1,i} \, a_i^i \\
m_{i+1}^{i+1} = R_{i+1,i} \, m_i^i
\end{cases}
\tag{9}
$$

where $R_{i+1,i} = R_{i+1,i}(q_{i+1})$ and $\dot{\theta}_{i+1} = \dot{q}_{i+1}$ when the relative motion of two consecutive frames depends on a revolute joint J_{i+1} in between, following the Denavit-Hartenberg parametrization (in this case, z_i is the $i-th$ joint axis), while $R_{i+1,i}$ is constant and $\dot{\theta}_{i+1} = 0$ otherwise.

The goal is to write the relationship between the measured variables in the frame $\{S_R\}$ of the scapula IMU and those in the frame $\{A_R\}$ attached to the arm IMU using the state variables. To do this, we first define the generic vector $\xi_n = \left[\omega_n^n, a_n^n, m_n^n \right]^T \in \mathbb{R}^9$, which contains all the variables associated to the n-*th* IMU in its frame $\{n\}$.

Choosing as measures $y = \xi_{A_R}$, i.e., the IMU measurements after the processing described in Section 3.1, the measurement model depends only on the state and on the output noise and results in:

$$
\begin{cases}
h = h(q, \dot{q}, \xi_{S_R}, \nu_S) \\
y = \xi_{A_R} + \nu_A
\end{cases}
\tag{10}
$$

The computation of h for the acceleration and magnetic field components is based on the simple relations $a_{A_R} = R_{A_R,S_R} a_{S_R}$ and $m_{A_R} = R_{A_R,S_R} m_{S_R}$, where the transformation R_{A_R,S_R} corresponds to:

$$
R_{A_R,S_R} = R_{C_A} \cdot R_q(q_1, q_2, q_3) \cdot R_{C_S}
\tag{11}
$$

where $R_q(q_1, q_2, q_3)$ is the rotation matrix between the Denavit-Hartenberg (D-H) frames, while R_{C_A} and R_{C_S} are the calibration rotation matrices obtained through the calibration procedure of Section 3.2. So, the acceleration and magnetic components of h depend only on q and ξ_{S_R}. The relation between the angular velocities ω_{S_R} and ω_{A_R} can be obtained by following the procedure in (9) from the first frame to the last one; in this case, the output function also depends on $\dot{q}$.

The magnetometer raw data are calibrated through the procedure described in Section 3.1. However, this step does not remove the disturbances that may affect the magnetic sensors, so we modified our UKF to increase the magnetometer noise to weigh this contribution less if a magnetic disturbance is acting on the sensor itself, as done in [37]. Indeed, if the norm of the magnetic field m does not fall within a certain range with respect to the normalized value $m_{norm} = 1$, we sensibly increase the noise variance of magnetometer measurements inside the output noise covariance matrix R of the UKF. In other terms, the magnetometer noise components σ_m^2 inside the matrix R were chosen as:

$$
\sigma_m^2 = f(||m|| - 1) + \sigma_{const}^2,
\tag{12}
$$

where $f(\cdot)$ is a function that depends linearly (or exponentially) on the difference $||m|| - 1$ through a parameter k (in our case, $f(||m|| - 1) = k(||m|| - 1)$, with $k = 10$).

Hence, the UKF allows us to estimate the shoulder joint angles q, leveraging on the inertial and magnetic field measures of the IMUs.

3. Experimental Setup

The goal of this experimental setup is to gather a set of data to validate both the UKF for the measurement of shoulder joint angles and the MVE to estimate the missing measurements for biomechanical assessment of the human arm.

We asked 9 able-bodied subjects (6 male and 3 female, age 28.2 ± 2.7, all right-handed) to perform the 30 tasks of daily living described in the SoftPro protocol [38]. Each of these

tasks was repeated three times for a total of 90 movements per subject. Participants did not have any physical limitations that could have affected the experimental outcomes. They gave their informed consent to participate. The procedures were approved by the Committee on Bioethics of the University of Pisa (Review No. 30/2020) in accordance with the Declaration of Helsinki. The pose in between movements consisted in resting the right hand flat on the table. Since these 90 movements were recorded in one shot, they were shuffled before being instructed to the subjects, to obtain an homogeneous dataset, not influenced by muscular fatigue.

The kinematic data were recorded with LSM9DS1 inertial sensors embedded in Arduino Nano 33 BLE boards and connected to a computer through serial communication at a sample rate of 120 Hz. The muscular data were recorded with the Delsys Bagnoli EMG System with a sampling frequency of 2400 Hz. The EMG placement followed SE-NIAM guidelines to minimize the cross-talk phenomen between near muscles is the same as the one adopted in the MHH dataset [38]. The EMG signals and the IMU data were recorded through a custom routine which guaranteed the synchronization between them. To validate the Kalman Filter results, we employed as a ground truth the Xsens MTw Awinda wearable system, which returns the upper-body posture of the subject. The kinematic data were recorded at the Xsens maximum sample rate of 60 Hz. To synchronize the Xsens data, collected via proprietary software, with the EMG and IMU signals, we performed Dynamic Time Warping (DTW) [39]. The whole sensor setup is shown in Figure 4.

3.1. IMU Processing

Before using the IMU data, it is important to remove constant biases that affect gyroscopes and accelerometers is important. An example of a debiasing routine can be found in [40]. The Arduino Nano 33 BLE boards, which were used for our work, directly provide the acceleration normalized with respect to the gravity acceleration $g = 9.81$ m/s^2.

Regarding the magnetic measures, the magnetometer raw data $^B m_r$ in the sensor frame $\{B\}$ lie on an ellipsoid manifold, as demonstrated in [41]. In the same work, to translate the raw data to the origin of the sensor frame and map them onto the unitary sphere, a Maximum Likelihood Estimator is used to determine the magnetometer optimal calibration parameters: a $SE(3)$ transformation matrix to align the ellipsoid axes with a calibration frame $\{C\}$ and center it on its origin, and a scaling matrix to stretch the ellipsoid on the unitary sphere. After this mapping, a second step allows us to find the optimal rotation matrix that minimizes the error between the data mapped on the unitary sphere $^C m$ and the original raw data $^B m_r$.

From a practical point of view, these calibration parameters can be determined with an initial data acquisition, during which the IMU should be rotated in as many configurations as possible. In this way, the shape of the ellipsoid can be better defined, avoiding sampling a small surface of the ellipsoid, for which the measurement noise can badly affect the parameter extraction.

3.2. IMU Frames Calibration

Prior to the estimation phase, it is necessary to evaluate the effective orientation of each sensor X attached to the body, i.e., to identify the rotation matrices between the sensor frames $\{S_R\}$ and $\{A_R\}$ and the first/last Denavit-Hartenberg frames, respectively.

In this section, we briefly introduce the approach used in our work and we direct the interested reader to [42] for more details. The procedure consists of a two-phase data acquisition: the first part is performed with the subject standing still with the arms straight along the body (N-pose); and in the second part, the subject is asked to slightly bend forward with their arm fixed to their body. These data return two readings of gravity acceleration in two different poses that are used in a series of cross products to define the calibration matrix.
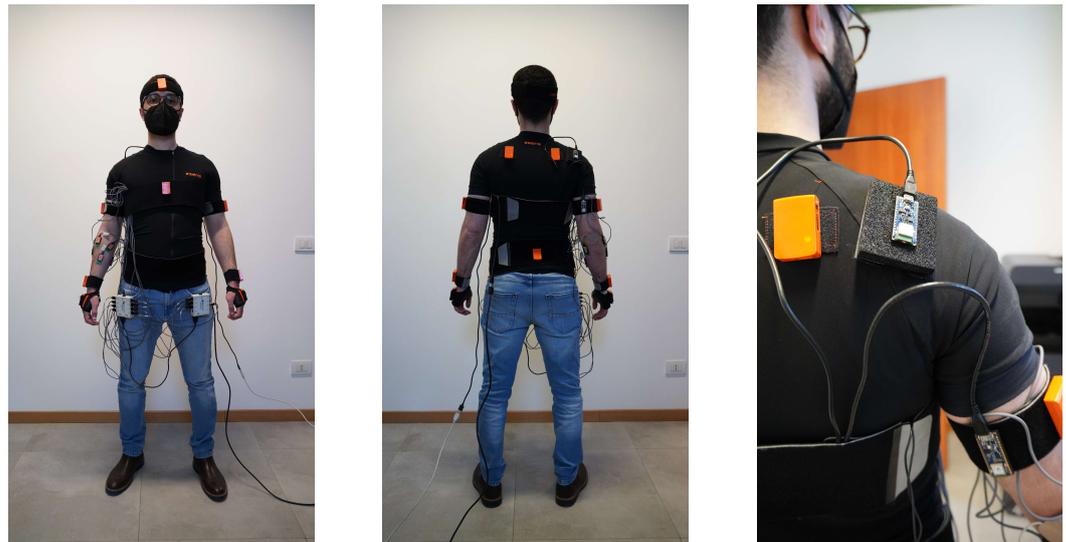
(**a**) Front view   (**b**) Back view   (**c**) IMU positioning

Figure 4. Different views of the complete sensor setup (including the ground truth sensors) used during the experimental phase. The full-body view of the system — composed by the Delsys Bagnoli EMG system (Delsys Inc., Natick, MA, USA), the Xsens MTw Awinda (Movella Inc., Henderson, NV, USA) wearable system and the two LSM9DS1 inertial sensors embedded in Arduino Nano 33 BLE boards (Arduino S.r.l., Monza, Italy) — is shown in (**a**,**b**). A detail of the IMUs positioning is depicted in (**c**).

3.3. EMG Processing

Surface EMG signals can be affected by different sources of noise (relative motion of soft tissues, bad mechanical or electrical connections, cross-talking between different muscles, etc...). Several works in literature provide solutions to this problem [43,44]. For our application, we took inspiration from [45] and we implemented the following filtering steps: (1) a first order low-pass Butterworth filter with a cutoff frequency of 500 Hz to reduce the high-frequency noise; (2) a first order high-pass Butterworth filter with a cutoff of 20 Hz, which allows us to remove the constant and slowly-changing behaviors; (3) the rectification of the filtered signal; and (4) another first order low-pass Butterworth filter, with a cutoff frequency of 1 Hz, for the extraction of the signal envelope.

3.4. From XSENS Quaternions to Joint Angles

For each link l of the arm kinematic chain, the XSENS system returns as output the quaternion Q_l, which expresses the orientation between the frame of the link and the system world frame. So, given the quaternions Q_s, Q_a and Q_f of the shoulder, arm and forearm respectively, we estimated the shoulder joint angles q_1, q_2 and q_3 and the elbow angles q_4 and q_5 through an Unscented Kalman Filter. Indeed, we can model the dynamics of the i-*th* joint angle as a random walk with Gaussian white noise w_{q_i}:

$$q_i(k+1) = q_i(k) + w_{q_i} \tag{13}$$

Then, we can use as measures y_1 for the estimation of the shoulder joints the orientation between the shoulder and arm link $y_1 = Q_{sa} = Q_s^* \otimes Q_a$, where $\otimes$ represents the quaternion product. Similarly, we can express the orientation between the arm and the forearm as $y_2 = Q_{af} = Q_a^* \otimes Q_f$ and use it as the second block of the output vector. So, the related output functions can be described as:

$$\begin{cases} h_1 = [Q_{01}(q_1) \otimes Q_{12}(q_2)] \otimes Q_{23}(q_3) \\ h_2 = Q_{34}(q_4) \otimes Q_{45}(q_5) \end{cases} \tag{14}$$

where the generic quaternion $Q_{i,i+1}$ express the orientation between two subsequent Denavit-Hartenberg frames through joint q_i.

4. Results

4.1. UKF Validation

To assess the UKF performance, three different metrics were used: the Root Mean Square (RMS) error between joint evolution estimated and the ones of the Xsens, used as ground truth; the Normalized Root Mean Square (NRMS) obtained by normalizing the RMS error with respect to the maximum range reached by each joint and the correlation index between the two signals (the UKF one and the ground truth) to evaluate their similarity in terms of temporal evolution.

Regarding the RMS, we reached a median value of around 10° (NRMS around 10%), with a performance comparable with other similar solutions presented in the literature [46–48], with an RMS error median between 5.2 to 7.9° in Slade et al., and between 4.95 to 7.03° in Peppoloni et al. The similarity between the estimated joint trajectories and the reference ones is also high, since it is about 0.93 for all the angles. In Table 1 the detailed results of these three metrics are reported, in terms of median and interquartile range, for each shoulder joint angle.

Table 1. UKF validation with respect to the Xsens system for shoulder joints estimation. In each column, RMS, Normalized RMS and correlation coefficient are reported in terms of the median and half of the interquartile range.

	RMS Error	NRMS Error	Correlation
q_1	$10.9 \pm 4.6°$	$11.27 \pm 4.72\%$	0.906 ± 0.084
q_2	$6.49 \pm 1.45°$	$6.93 \pm 1.525\%$	0.956 ± 0.028
q_3	$11.1 \pm 3.85°$	$11.01 \pm 3.79\%$	0.930 ± 0.07

4.2. MVE Validation

To evaluate the goodness of estimation performed by MVE, we computed the RMS error (RMSE) and NRMS error (NRMSE) comparing it with the ground truth value recorded during tasks execution. In Figure 5 the NRMSE between the real signal and the output of the MVE for each DoF is reported in terms of the median and interquartile range. The measured DoFs are represented in blue, while the estimated ones are in red. For the kinematic part, the NRMS error on the measured joints is about 2.4%. We can notice, as expected, a higher error for the estimated joints with respect to the measured ones, with a median around 8.5%. However, the error level is comparable with the one reached in other solutions presented in the literature [36], with the advantage of a lower number of used sensor elements. For the muscular side, the normalized error level achieved is even lower (maximum median NRMSE just above 4%).

In terms of RMSE, it reaches $17.1 \pm 4.97°$ for the non-measured joint angles, while for the muscles it is 0.003 ± 0.002 mV (values expressed in median $\pm$ interquartile range). This result, compared to the one reported in [34] ($2.18 \pm 1.32°$ for the joints and 0.003 ± 0.002 mV for the muscles), can be considered sufficiently good, as this joint angle choice was not the optimal one found in [34] and referred to a selection of individual DoFs, but represents an approximation that fulfills the requirement of the minimum number of sensors required for an effective implementation of the measurements. Furthermore, in [34], the kinematic measurements considered for the analysis were provided by a ground truth optical system, while in our case we used the information measured by the IMU-based system we developed - which intrinsically comes with an estimation error, although comparable with or less than the one of the other related works in the literature. An example of a random estimated movement is presented in Figure 6. The not measured DoFs are marked with a star (*). These graphs confirm the results obtained in terms of RMS error.
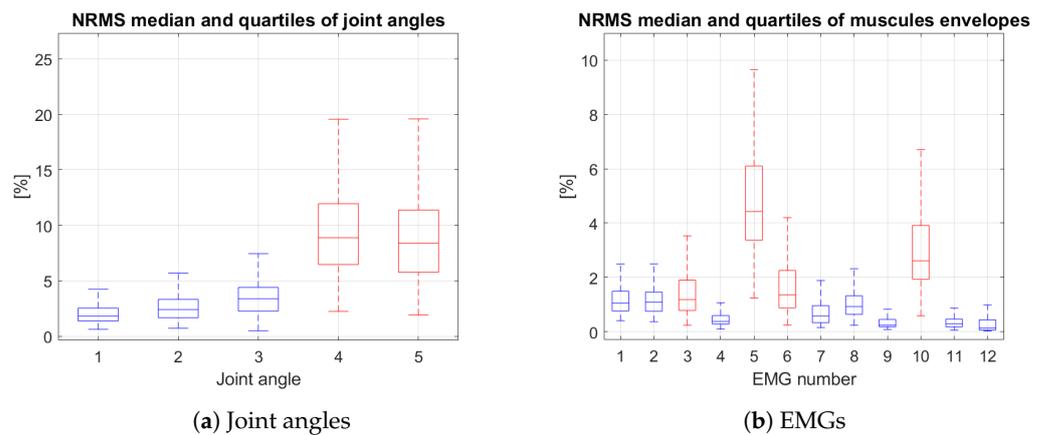
(a) Joint angles   **(b)** EMGs

Figure 5. Normalized RMS Error computed for each DoF (measured DoFs in blue, non-measured DoFs in red).

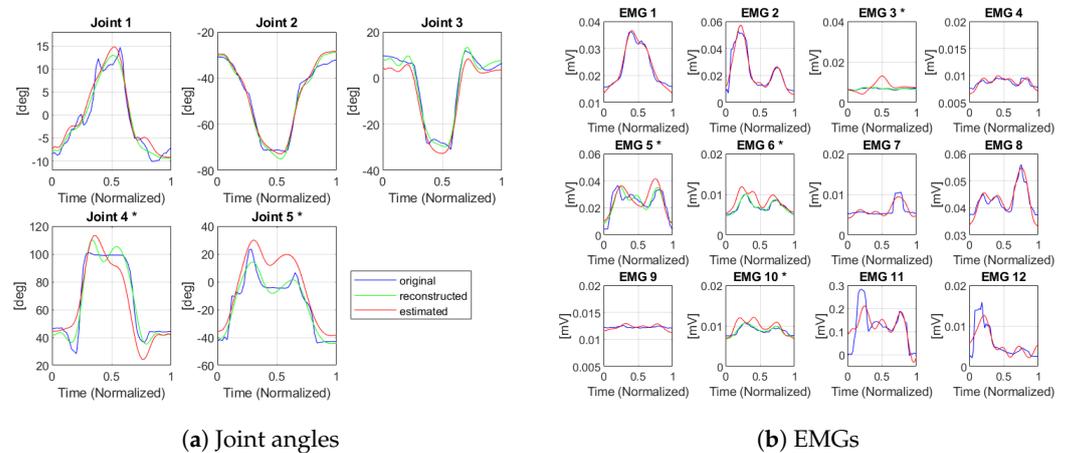**(a)** Joint angles   **(b)** EMGs

Figure 6. Example of MVE on a movement of the test dataset (in blue: reference movement; in green: movement reconstruction with fPCs; in red: movement obtained through MVE); * = non−measured DoFs.

5. Conclusions

The topic of human-robot interaction and collaboration, as well as monitoring the human musculoskeletal state in working environments, has gained increasing attention in recent years. In particular, the assessment of the musculoskeletal state could bring many benefits in terms of improving working conditions and preventing work-related disorders.

In this paper, we present a technological solution that relies on a reduced number of wearable sensing units (IMUs and sEMGs) and provides an estimation of the whole musculoskeletal state.

To do this, we developed an under-sensorized wearable system that exploits the Minimum Variance Estimation approach to assess the bio-mechanical state of the human arm. Additionally, an Unscented Kalman Filter was implemented to directly obtain the joint angle trajectories from the IMUs measurements. This setup was extensively tested through the collection of a new dataset of daily living activities. The obtained results are promising, as they show an average normalized error of 8.5% on the non-measured joints and of 2.5% on the non-measured EMGs. Our system allows an accurate state monitoring, with a reduced number of sensors, thus increasing wearability and reducing discomfort.

Our outcomes can pave the path toward unobtrusive wearable monitoring of multi-modal quantities. First, our theoretical framework allows us to overcome the limitations of data-driven methods that rely on the usage of large training datasets that can be used to complement scarce sensory information. Of note, such a theoretical framework was already presented in our previous publication [34]. Second, we provided, for the first time,

an implementation of our optimal design, showing that, with a reduced set of optimally placed sensors, we can reconstruct the whole musculoskeletal state of the upper limb. This under-sensorized implementation leads to the reduction of the number of sensors, enhancing the overall system wearability. While this is already a good achievement for the monitoring of the upper limb, our implementation can pave the path toward whole-body multi-modal sensing, where ergonomics and economic constraints pose even more strict constraints on the number, and quality, of sensors in use.

Starting from these results, the next step will be to compare this approach with a fully data-driven approach (e.g., Deep Generative Adversarial Network [49]) to evaluate the performance of our MVE-based solution with respect to the ones obtained by deep learning techniques, and eventually propose hybrid approaches. Another interesting path to explore would be to find a way to use this setup online, as the functional decomposition requires a movement to be recorded in advance. In the future, we will investigate other techniques for the fusion of IMU and EMG data—and compare and integrate them with our approach also targeting action recognition. It will also be interesting to study zero crossing/time-frequency domain for gesture recognition and HRI [50,51].

Finally, these methods could be extended to the entire human body and therefore assess the entire skeletal and muscular state of a person in different application contexts, such as rehabilitation and human-robot collaboration.

Institutional Review Board Statement: The procedures were approved by the Committee on Bioethics of the University of Pisa (Review No. 30/2020) in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data are not publicly available due to privacy reasons.

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

References

1. Bai, L.; Pepper, M.G.; Yan, Y.; Phillips, M.; Sakel, M. Low Cost Inertial Sensors for the Motion Tracking and Orientation Estimation of Human Upper Limbs in Neurological Rehabilitation. *IEEE Access* **2020**, *8*, 54254–54268. [CrossRef]
2. Worsey, M.T.; Espinosa, H.G.; Shepherd, J.B.; Thiel, D.V. Inertial sensors for performance analysis in combat sports: A systematic review. *Sports* **2019**, *7*, 28. [CrossRef] [PubMed]
3. Rana, M.; Mittal, V. Wearable sensors for real-time kinematics analysis in sports: A review. *IEEE Sens. J.* **2020**, *21*, 1187–1207. [CrossRef]
4. Lasota, P.A.; Fong, T.; Shah, J.A. A survey of methods for safe human-robot interaction. *Found. Trends Robot.* **2017**, *5*, 261–349. [CrossRef]
5. Maurice, P.; Malaisé, A.; Amiot, C.; Paris, N.; Richard, G.J.; Rochel, O.; Ivaldi, S. Human movement and ergonomics: An industry-oriented dataset for collaborative robotics. *Int. J. Robot. Res.* **2019**, *38*, 1529–1537. [CrossRef]
6. Ranavolo, A.; Ajoudani, A.; Cherubini, A.; Bianchi, M.; Fritzsche, L.; Iavicoli, S.; Sartori, M.; Silvetti, A.; Vanderborght, B.; Varrecchia, T.; et al. The Sensor-Based Biomechanical Risk Assessment at the Base of the Need for Revising of Standards for Human Ergonomics. *Sensors* **2020**, *20*, 5750. [CrossRef] [PubMed]

7. Shafti, A.; Ataka, A.; Lazpita, B.U.; Shiva, A.; Wurdemann, H.A.; Althoefer, K. Real-time robot-assisted ergonomics. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1975–1981.

8. Cop, C.P.; Cavallo, G.; van't Veld, R.C.; Koopman, B.F.; Lataire, J.; Schouten, A.C.; Sartori, M. Unifying system identification and biomechanical formulations for the estimation of muscle, tendon and joint stiffness during human movement. *Prog. Biomed. Eng.* **2021**, *3*, 033002. [CrossRef]

9. Holzbaur, K.R.; Murray, W.M.; Gold, G.E.; Delp, S.L. Upper limb muscle volumes in adult subjects. *J. Biomech.* **2007**, *40*, 742–749. [CrossRef]

10. Sengupta, A.; Cao, S. mmPose-NLP: A Natural Language Processing Approach to Precise Skeletal Pose Estimation Using mmWave Radars. In *IEEE Transactions on Neural Networks and Learning Systems*; IEEE: Piscataway, NJ, USA, 2022.

11. Zhao, M.; Li, T.; Abu Alsheikh, M.; Tian, Y.; Zhao, H.; Torralba, A.; Katabi, D. Through-wall human pose estimation using radio signals. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7356–7365.

12. Filippeschi, A.; Schmitz, N.; Miezal, M.; Bleser, G.; Ruffaldi, E.; Stricker, D. Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion. *Sensors* **2017**, *17*, 1257. [CrossRef]

13. Kok, M.; Hol, J.D.; Schön, T.B. Using Inertial Sensors for Position and Orientation Estimation. *arXiv* **2017**, arXiv:1704.06053.

14. Rampichini, S.; Vieira, T.M.; Castiglioni, P.; Merati, G. Complexity analysis of surface electromyography for assessing the myoelectric manifestation of muscle fatigue: A review. *Entropy* **2020**, *22*, 529. [CrossRef] [PubMed]

15. Rocha, V.d.A.; do Carmo, J.C.; Nascimento, F.A.d.O. Weighted-cumulated S-EMG muscle fatigue estimator. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1854–1862. [CrossRef] [PubMed]

16. Jebelli, H.; Lee, S. Feasibility of wearable electromyography (EMG) to assess construction workers' muscle fatigue. In *Advances in Informatics and Computing in Civil and Construction Engineering*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 181–187.

17. Woodward, R.B.; Shefelbine, S.J.; Vaidyanathan, R. Pervasive monitoring of motion and muscle activation: Inertial and mechanomyography fusion. *IEEE/ASME Trans. Mechatron.* **2017**, *22*, 2022–2033. [CrossRef]

18. Pacchierotti, C.; Sinclair, S.; Solazzi, M.; Frisoli, A.; Hayward, V.; Prattichizzo, D. Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives. *IEEE Trans. Haptics* **2017**, *10*, 580–600. [CrossRef]

19. Latash, M.L. One more time about motor (and non-motor) synergies. *Exp. Brain Res.* **2021**, *239*, 2951–2967. [CrossRef]

20. Zatsiorsky, V.M.; Latash, M.L. Prehension synergies. *Exerc. Sport Sci. Rev.* **2004**, *32*, 75. [CrossRef]

21. Latash, M.L. *Neurophysiological Basis of Movement*; Human Kinetics: Champaign, IL, USA, 2008.

22. Scano, A.; Dardari, L.; Molteni, F.; Giberti, H.; Tosatti, L.M.; d'Avella, A. A comprehensive spatial mapping of muscle synergies in highly variable upper-limb movements of healthy subjects. *Front. Physiol.* **2019**, *10*, 1231. [CrossRef]

23. Averta, G.; Valenza, G.; Catrambone, V.; Barontini, F.; Scilingo, E.P.; Bicchi, A.; Bianchi, M. On the time-invariance properties of upper limb synergies. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 1397–1406. [CrossRef]

24. Van Criekinge, T.; Vermeulen, J.; Wagemans, K.; Schröder, J.; Embrechts, E.; Truijen, S.; Hallemans, A.; Saeys, W. Lower limb muscle synergies during walking after stroke: A systematic review. *Disabil. Rehabil.* **2020**, *42*, 2836–2845. [CrossRef]

25. Chvatal, S.A.; Ting, L.H. Voluntary and reactive recruitment of locomotor muscle synergies during perturbed walking. *J. Neurosci.* **2012**, *32*, 12237–12250. [CrossRef]

26. Catalano, M.G.; Grioli, G.; Farnioli, E.; Serio, A.; Piazza, C.; Bicchi, A. Adaptive synergies for the design and control of the Pisa/IIT SoftHand. *Int. J. Robot. Res.* **2014**, *33*, 768–782. [CrossRef]

27. Ciocarlie, M.T.; Allen, P.K. Hand posture subspaces for dexterous robotic grasping. *Int. J. Robot. Res.* **2009**, *28*, 851–867. [CrossRef]

28. Averta, G.; Della Santina, C.; Valenza, G.; Bicchi, A.; Bianchi, M. Exploiting upper-limb functional principal components for human-like motion generation of anthropomorphic robots. *J. Neuroeng. Rehabil.* **2020**, *17*, 1–15. [CrossRef] [PubMed]

29. Ficuciello, F.; Palli, G.; Melchiorri, C.; Siciliano, B. Experimental evaluation of postural synergies during reach to grasp with the UB Hand IV. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 1775–1780.

30. Bianchi, M.; Salaris, P.; Bicchi, A. Synergy-based hand pose sensing: Reconstruction enhancement. *Int. J. Robot. Res.* **2013**, *32*, 396–406. [CrossRef]

31. Bianchi, M.; Salaris, P.; Bicchi, A. Synergy-based hand pose sensing: Optimal glove design. *Int. J. Robot. Res.* **2013**, *32*, 407–424. [CrossRef]

32. Ciotti, S.; Battaglia, E.; Carbonaro, N.; Bicchi, A.; Tognetti, A.; Bianchi, M. A synergy-based optimally designed sensing glove for functional grasp recognition. *Sensors* **2016**, *16*, 811. [CrossRef] [PubMed]

33. Kuismin, M.O.; Sillanpää, M.J. Estimation of covariance and precision matrix, network structure, and a view toward systems biology. *Wiley Interdiscip. Rev. Comput. Stat.* **2017**, *9*, e1415. [CrossRef]

34. Averta, G.; Iuculano, M.; Salaris, P.; Bianchi, M. Optimal Reconstruction of Human Motion From Scarce Multimodal Data. *IEEE Trans.-Hum.-Mach. Syst.* **2022**, *52*, 833–842. [CrossRef]

35. Ramsay, J.O.; Hooker, G.; Graves, S. *Functional Data Analysis with R and MATLAB*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.

36. Peppoloni, L.; Filippeschi, A.; Ruffaldi, E.; Avizzano, C.A. A novel 7 degrees of freedom model for upper limb kinematic reconstruction based on wearable sensors. In Proceedings of the 2013 IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY); IEEE: Piscataway, NJ, USA, 2013; pp. 105–110.

37. Saito, A.; Kizawa, S.; Kobayashi, Y.; Miyawaki, K. Pose estimation by extended Kalman filter using noise covariance matrices based on sensor output. *Robomech J.* **2020**, *7*, 1–11. [CrossRef]

38. Averta, G.; Barontini, F.; Catrambone, V.; Haddadin, S.; Handjaras, G.; Held, J.P.; Hu, T.; Jakubowitz, E.; Kanzler, C.M.; Kühn, J.; et al. U-Limb: A multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions. *GigaScience* **2021**, *10*, giab043. [CrossRef]

39. Müller, M. Dynamic time warping. In *Information Retrieval for Music and Motion*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 69–84.

40. Tedaldi, D.; Pretto, A.; Menegatti, E. A robust and easy to implement method for IMU calibration without external equipments. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA); IEEE: Piscataway, NJ, USA, 2014; pp. 3042–3049.

41. Vasconcelos, J.F.; Elkaim, G.; Silvestre, C.; Oliveira, P.; Cardeira, B. Geometric approach to strapdown magnetometer calibration in sensor frame. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 1293–1306. [CrossRef]

42. Bleser, G.; Hendeby, G.; Miezal, M. Using egocentric vision to achieve robust inertial body tracking under magnetic disturbances. In Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality; IEEE: Piscataway, NJ, USA, 2011; pp. 103–109.

43. Farago, E.; Macisaac, D.; Suk, M.; Chan, A.D.C. A Review of Techniques for Surface Electromyography Signal Quality Analysis. *IEEE Rev. Biomed. Eng.* **2022**, *16*, 472–486. [CrossRef]

44. Zhao, Y.; Zhang, S.; Yu, T.; Zhang, Y.; Ye, G.; Cui, H.; He, C.; Jiang, W.; Zhai, Y.; Lu, C.; et al. Ultra-conformal skin electrodes with synergistically enhanced conductivity for long-time and low-motion artifact epidermal electrophysiology. *Nat. Commun.* **2021**, *12*, 4880. [CrossRef]

45. Potvin, J.; Brown, S. Less is more: High pass filtering, to remove up to 99% of the surface EMG signal power, improves EMG-based biceps brachii muscle force estimates. *J. Electromyogr. Kinesiol.* **2004**, *14*, 389–399. [CrossRef]

46. El-Gohary, M.; McNames, J. Human joint angle estimation with inertial sensors and validation with a robot arm. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1759–1767. [CrossRef]

47. Slade, P.; Habib, A.; Hicks, J.L.; Delp, S.L. An open-source and wearable system for measuring 3D human motion in real-time. *IEEE Trans. Biomed. Eng.* **2021**, *69*, 678–688. [CrossRef] [PubMed]

48. Alizadegan, A.; Behzadipour, S. Shoulder and elbow joint angle estimation for upper limb rehabilitation tasks using low-cost inertial and optical sensors. *J. Mech. Med. Biol.* **2017**, *17*, 1750031. [CrossRef]

49. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [CrossRef]

50. Liu, H.; Hartmann, Y.; Schultz, T. A Practical Wearable Sensor-based Human Activity Recognition Research Pipeline. In Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies, Online, 9–11 February 2022; pp. 847–856.

51. Hartmann, Y.; Liu, H.; Schultz, T. Interactive and Interpretable Online Human Activity Recognition. In Proceedings of the 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Pisa, Italy, 21–25 March 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 109–111.