Feature Space Exploration for Motion Classification Based on Multi-Modal Sensor Data for Lower Limb Exoskeletons

Tilman Daab, Isabel Patzer, Ralf Mikut and Tamim Asfour

***Abstract*— In this paper, we address the problem of finding a minimal multi-modal sensor setup for motion classification in lower limb exoskeleton applications while maintaining the classification performance. We present an approach for a systematic exploration of the feature space and feature space dimensionality reduction for motion recognition using Hidden Markov Models (HMMs). We evaluated our approach using IMU and force sensor data with 10 subjects performing 14 different daily activities. We perform a dimensionality reduction on sensor feature level with single- and multi-subjects and we explore the feature space using fine-grained features such as the force value of a single direction. Additionally, we investigate the influence of physical characteristics on the classification quality. Our results show that a subject specific and general reduction of the sensors is possible while still achieving the same classification performance.**

I. Introduction

Human action recognition has been a large research field over the last years. It covers many different topics such as hand gesture recognition, whole body motion recognition, semantic segmentation and imitation learning. Hence, the recognition can be based on different types of input data such as visual data ([1], [2]), data of wearable sensors ([3], [4]) or a combination of such approaches ([5]). In the context of human action recognition for exoskeletons often wearable sensors are used. They have the advantage of allowing the exoskeleton wearer to move independently and without being restricted to a certain area. One important requirement for commercial wearable devices, such as exoskeletons, prosthetics and orthotics, is the reduction of the number of sensors due to several factors such as costs, limited computing resources and energy consumption. Therefore, it is essential to identify a minimal set of sensors, which still allows a correct and robust motion classification.

In our previous work, we introduced an unilateral lower limb exoskeleton (KIT-EXO-1) with a force-based interface to the human leg and two active Degrees of Freedom (DoF), see [6]. In [7], we presented a motion classification system based on Hidden Markov Models (HMMs) and evaluated it using a new unilateral, passive lower limb exoskeleton for the left leg, shown in Figure 1. This exoskeleton is equipped with a sensor system consisting of 3 Inertial Measurement Units (IMUs), one on each segment and seven 3D force sensors (FS). These force sensors are arranged in a way that corresponds to the main muscles being involved in human locomotion. This arrangement allows the measurement of interaction forces between the exoskeleton and the human lower limb. We evaluated the online classification performance with multi-subject data using a sliding-window approach. In particular, we investigated the generalization performance and latency. Based on these results, we presented a single- vs. multi-subject evaluation, a detailed analysis of the classification performance of each single motion type and investigated a minimal sensor setup for lower limb exoskeletons based on a brute-force approach in [8].

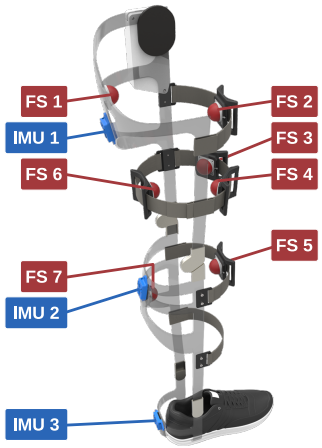


Fig. 1: Passive exoskeleton with seven 3D-force (red) and three IMU sensors (blue). The numbers correspond to the sensor labels.

In this paper, we introduce an approach for a systematic exploration of the feature space for motion classification, which is based on dimensionality reduction techniques for the recognition of whole-body human motions described in [9]. We investigate how the presented techniques can be applied to motion classification in exoskeletons based on multi-modal sensory data. Our analyses cover two main aspects: the sensory-based dimensionality reduction and feature-based dimensionality reduction. Thereby, we address the question of a minimal set of sensors (respectively features) required to achieve a certain motion classification accuracy. Furthermore, we investigate how the body characteristics of different subjects influence the classification performance. To allow a representative comparison, we use the same data used in our previous work described in [7] in our evaluation.

The paper is organized as follows. Section II provides a brief overview of sensor setups, machine learning methods

This work has been supported by the German Federal Ministry of Education and Research (BMBF) under the project INOPRO (16SV7665).

T. Daab, I. Patzer and T. Asfour are with the Institute for Anthropomatics and Robotics. T. Daab is a scholar of the Heinrich-Böll-Stiftung. R. Mikut is with Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology, Germany.
{isabel.patzer, ralf.mikut, asfour}@kit.edu; tilman.daab@student.kit.edu

and their combinations in the context of human action recognition as well as different feature reduction approaches. In Section III, we give a short recap of our previous works regarding the passive lower limb exoskeleton, the underlying data and the HMM-based motion classification approach. Furthermore, we want to introduce a baseline for motion classification conditions in context of exoskeletons. Section IV describes our approach to systematically explore the feature space. Section V describes the evaluation results while Section VI provides a conclusion and an outlook.

II. Related Work

Classification of human motions is used in the field of industry to evaluate the ergonomic of the workers [10] and to track time of single working steps [11]. A further field of application is using these sensors as an interface to the human body for the control of exoskeletons ([6], [7], [8], [12]). In the context of motion classification different sensors can be used, such as depth cameras, Inertial Measurement Units (IMUs), force sensors, EMG- and EEG-sensors. Based on the application, a suitable sensor setup has to be chosen. In the context of exoskeletons often wearable sensors such as EMG-, EEG- or mechanical sensors are used [12]. Tsai et al. [13] analyzed motion patterns in multi-channel EMG signals for the control of an upper arm exoskeleton robot. Several lower limb exoskeletons are controlled via EEG-based interfaces ([14], [15], [16]). A combination of EEG and surface EMG signals is presented in [17] where the H2 exoskeleton is controlled by identifying signal patterns using Artificial Neuronal Networks and Support Vector Machines (SVMs). A disadvantage of using EEG or EMG sensors is the classification accuracy due to the subject specific nature of the signal patterns and the temporal change of the signal ([18], [19], [20]). In other applications, sensors such as IMUs or torque sensors, are often used for motion classification, e.g. it is especially common in industrial environments to place inertial sensors on the human body ([21], [22], [23]).

Jang et al. [12] use three IMUs in combination with two joint angle encoders for motion classification in the context of exoskeletons. Based on SVMs they distinguish walking on flat ground, on a ramp and on stairs (upwards and downwards). Attal et al. [21] compare different machine learning approaches, such as SVMs and Hidden Markov Models (HMMs), by using input from three inertial sensors placed on the left ankle, the right thigh and the chest. Malaisé et al. use a wearable motion tracking suit (MVN Link suit Xsens) with 17 integrated IMUs and an additional sensorized glove (Emphasis Telematics) for motion classification via HMMs in an industrial environment ([10], [24]).

Some works address the problem of reducing the number of sensors and features to scale down manufacturing costs and computational effort while keeping the classification and prediction accuracy. One strategy to achieve this is to reduce the number of dimensions by aggregation, e.g. via Principal Component Analysis (PCA). The amount of the input data remains the same but is reduced to a low dimensional description. Another possibility is to use algorithms for selective dimensionality reduction. To evaluate the generated features, there are three common methods: embedded methods, filter-based methods and wrapper-based methods. Malaisé et al. [10] compared PCA, a filter-based, and a wrapper-based method for their HMM-based motion classification. In their analysis they used the data of 13 subjects and showed that a wrapper-based method performed best compared to the other two. Their wrapper-based method is based on our previous work [9] where we conducted a feature space dimensionality reduction for the recognition of whole-body human actions based on Hidden Markov Models.

III. Sensor System and Motion Classification

We provide a brief overview of our previous works ([7], [8]) on motion classification, system setup, data and methods which are used for the evaluations in Section V. For detailed information on these topics, we refer to [7] and [8].

A. Exoskeleton and Sensors

For this work, we use the data captured with our passive lower limb exoskeleton for the left leg (Figure 1). The frames of the exoskeleton cover the thigh, shank and foot. Two orthotic revolute joints connect these three components. The orientations and linear accelerations are measured on each limb segment (thigh, shank and foot) by a total of 3 Inertial Measurement Units (IMUs). The interaction forces between the exoskeleton and the wearer are measured by 7 individual 3D force sensors. The force sensors are placed over large muscles on the front and back of the thigh, as well as on the shank, which are mainly involved in walking motions.

B. Data

Our previously collected data described in [7] and [8] consists of 14 different motion activities, namely: *Walking Forward (WF), Walking Backward (WB), Turn Left (TL), Turn Right (TR), Sidesteps Right (SR), Sidesteps Left (SL), Going Upstairs (GU), Going Downstairs (GD), Going Downstairs Backwards (GDB), Lift Object (LO), Drop Object (DO), Stand Up (SU), Sit Down (SD)* and *Stand (ST)*. Each of the 10 subjects (5 male, 5 female) performed every motion 10 times. The timestamps of the recordings were unified, since the IMU data (80 Hz) and the force data (100 Hz) were recorded at a different frequency.

C. Motion Classification

Our evaluations and results are based on a Hidden Markov Model (HMM) multi-class classification approach. We trained one HMM for each motion, leading to 14 HMMs in total. We used a sliding window approach to enable an online application. Therefore, we split the data into windows of 100, 200 or 300 ms, depending on the evaluation. A new window starts every 10 ms. A currently tested window is assigned to the HMM with the highest log-likelihood.

We used the values of the force and IMU sensors as training input. The force feature vector consists of the 3D force data of every force sensor, resulting in a total of

21 values for all seven force sensors. The IMU feature vector contains the 3D linear acceleration, as well as the Roll-Pitch-Yaw angles for every segment. This leads to an 18 dimensional IMU feature vector (containing all 3 IMUs).

D. Baseline

Using wearable sensors for human action recognition can be evaluated under different aspects. One is the placement of the sensors on the human body or exoskeleton. Depending on the motions the system has to recognize, a wide range of sensor placements is possible. In the context of leg motions, often the foot, shank and thigh are used for the placement of the IMU sensors ([10], [12], [21]). It can be observed that using multiple sensors increases the classification performance compared to using single sensors ([8], [21], [25]). When using wearable sensors the latency (time between the activation of the movement by the user and the action being correctly classified) varies strongly on the application of the exoskeleton. When observing human motions in the context of industrial applications often windows with a width of more than 1 second are common ([26]), whereas in the context of fall prevention or human augmentation a very low latency is required ([26]). For fast adaptations to the human movements latencies between 100-300 ms are desirable ([27]). In [28], the authors suggest that a window size of 150-250 ms provides a good trade-off between the classification accuracy and latency. This is frequently realized by using overlapping windows [29].

A further important point is the classification performance of the system. It is desirable to achieve an accuracy near 100 % since the exoskeleton has to classify motions as error free as possible. Even achieving an accuracy of about 95 % would mean by one million steps, that roughly 50 thousand of these steps are wrongly classified and could therefore lead to increasing interaction forces between the human body and the exoskeleton which in turn lead to a decreased wearing comfort. One solution to overcome this problem is to implement further strategies which do not only take one single classification into account but take the entire locomotion task into consideration, e. g. to allow using knowledge about the previously recognized actions.

IV. Feature Selection Approach

For the feature selection approach in this paper, we use an adaptation of the wrapper-based method we developed in our previous work [9]. There, we conducted a feature space dimensionality reduction for the recognition of whole-body human motions based on Hidden Markov Models. Our adapted approach consists of the following steps:

1) Feature definition
2) Search among features
    a) Build feature sets of certain size
    b) Train and test classifier
    c) Identify feature sets for extension
    d) Repeat until maximum size
3) Evaluate different feature sets

In the *Feature definition* step the features are defined based on the chosen sensor setup. To define features, the measured values have to be processed and grouped. This can be done at different levels of detail.

Let us assume that Ω is the set of all sensor modalities. Each modality $\Gamma \in \Omega$ can have n sensors. Examples for modalities are Inertial Measurement Unit (IMU) sensors, force sensors or angle encoders. Each modality Γ is represented as a single capital letter, e.g. A or B. Each sensor of modality Γ is defined by the same feature representations $f_k(\Gamma)$. The sensor data can be represented at different levels of detail. The data type of modality Γ at detail level k is described by $f_k(\Gamma)$, $k \in \{1, 2, ..., m\}$. The most aggregated form of the sensor data is represented by $f_m(\Gamma)$. This combined representation describes hierarchical feature sets and can be subdivided into features of the form $f_k(\Gamma)$, $k \in \{1, 2, ..., m-1\}$. $f_1(\Gamma)$ describes thereby only scalar values. Based on the use case, the level of detail k can be chosen separately for each modality Γ.

The result of applying step *1 – Feature definition* to the exoskeleton data introduced in Section III-C is depicted in Table I. The data is represented by two different sensor modalities Γ, namely A which represents the IMU sensors and B which represents the force sensors. This leads to $\Omega = \{A, B\}$. A has 2 data types, the Euler angles (e_x, e_y, e_z) and the linear accelerations (a_x, a_y, a_z) in x, y, z direction. These sensor modalities can be used in combination $f_3(A) = ((e_x, e_y, e_z), (a_x, a_y, a_z))$ or can be split into smaller features, $f_2(A) = (e_x, e_y, e_z)$ and (a_x, a_y, a_z), or scalar values $f_1(A) = e_x$, e_y, e_z, a_x, a_y and a_z. The force sensor data B is represented by forces F_x, F_y, F_z in all three directions. $f_3(B)$ is empty since a further aggregation of the force values is not possible.

TABLE I: Step *1* applied to exoskeleton data.

Γ	A [IMU]	B [Force Sensor]
$f_1(\Gamma)$	e_1 e_2 e_3 a_x a_y a_z	F_x F_y F_z
$f_2(\Gamma)$	(e_1, e_2, e_3) (a_x, a_y, a_z)	(F_x, F_y, F_z)
$f_3(\Gamma)$	$((e_1, e_2, e_3), (a_x, a_y, a_z))$	-

A non greedy-search ([9]) among the previously defined features is then conducted in step *2 – Search among features*. In step *2a* feature sets of a certain size are built. We extend feature sets by one feature so that the dimensionality of the resulting feature set matches the number of the iteration. At iteration 1 the single different feature sets equal scalar values. In iteration 2 the feature set is extended by one additional feature. For more details on the feature selection approach we refer to [9].

In step *2b – Train and test classifier* the current feature set is trained and tested with the chosen classifier, e.g. a Hidden Markov Model or a Support Vector Machine. To allow comparing features sets, a performance criterion has to be selected, e.g. accuracy or F_1 score. Testing the feature sets of the current dimension yields a score for each of them.

In the next step *2c – Identify feature sets for extension* the best s feature sets of the current iteration are saved and are used for subsequent iterations. In step *2a*, such feature

sets are extended by one additional feature to form a new feature set. Extending s feature sets instead of just one reduces the risk of eliminating important feature subsets. The steps *2a–c* are repeated until the maximum dimension is reached. Afterwards the different feature sets can be evaluated, depending on the research question (step *3 – Evaluate different feature sets*).

For step *2b*, we use a Hidden Markov Model classifier as in [8] and [9]. The classification performance is assessed using the F_1 score. For training and testing, we perform a stratified 3-fold cross validation as used in [8] and [9]. The number of best features s is set to $s = 10$ ([9]).

V. Evaluations and Results

In this section the different evaluation methods and their results regarding the sensor-based and fine-grained dimensionality reduction are presented.

In our previous work [8] numerous feature sets differed only marginally in their performance (0.01 %–2.06 %). While selecting the best n feature sets for further evaluation might often be adequate, such procedure likely discard relevant feature sets in this case. To avoid this, we count the frequency of features in suitable feature sets. These are all sets with a score differing no more than 3.5 % from the best feature set of the same dimension. If this threshold rates all feature sets of this dimension as suitable, it is lowered to 1.0 %.

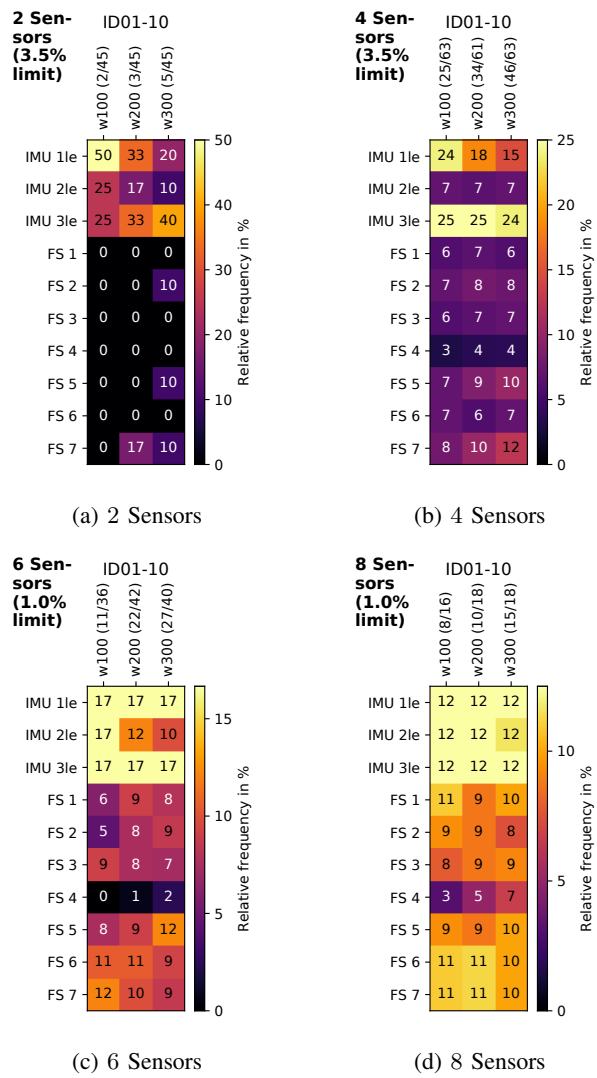
A. Sensor-based Reduction

For all subjects individually and in combination sensor-based tests were performed using the window sizes 100 ms, 200 ms and 300 ms to determine the suitability of sensors, the influence of the test subject and the influence of the window size. The sensor-wise suitability test corresponds to an analysis on a hardware level which means that at least 1 sensor and at most 10 sensors of the exoskeleton are used. Here, irrelevant sensors should be identified in order to reduce the required sensor setup. The used features equal $f_3(A)$ and $f_2(B)$. In consequence, the dimension of all evaluated feature sets is a multiple of three.

*1) **Multi-Subject**:* An excerpt of the evaluation of multi-subjects is depicted in Figure 2. The columns correspond to the window size, the rows to the used feature. *IMU-xle* describes the x-th sensor (see positions of the sensors in Figure 1) of the IMU sensors using all linear accelerations and Euler angles of this single sensor combined. *FS-y* corresponds to the y-th force sensor. The suitability threshold changes from 3.5% to 1.0% for 6 and 8 used sensors, since the results are otherwise too close to each other and all combinations would be taken into account. The numbers (a/b) following the window size describe the number of suitable feature sets a among all evaluated feature sets b in this round.

Figure 2b shows the evaluation results for combining four sensors using the data of all 10 subjects. 63 feature sets were evaluated for a window size of 100 ms (first column). 25 of these features were suitable, i.e. within the threshold of 3.5%. The colours of the heatmap describe the relative frequency of occurrence of the single sensors. While *IMU-1le* has a 24% share, *IMU-3le* has a 25% share. The other sensors appear between 3% and 8%. The further columns are read equivalently.

For combinations of up to 3 sensor representing features the IMU sensors are dominant. The force sensors rarely occur. Overall, *IMU-2le* clearly occurs less frequently than *IMU-1le* and *IMU-3le*, which usually occur with a similar frequency or a slight advantage for *IMU-1le*. As depicted in Figure 2a, *IMU-2le* is more relevant for small window sizes. Among the force sensors, *FS-7* occurs mostly. The occurrence of *IMU-2le* decreases in Figure 2b compared to Figure 2a. *IMU-1le* and *IMU-3le* remain dominant. The force sensors *FS-5* and *FS-7* occur slightly more frequently than the other force sensors.



(a) 2 Sensors (b) 4 Sensors

(c) 6 Sensors (d) 8 Sensors

Fig. 2: Heatmap for the multi-subject evaluations with different window sizes and different number of sensors.

For combinations of 6 sensors, the frequency of occurrence of *IMU-2le* compared to *IMU-1le* and *IMU-3le* is similar for a window size of 100 ms and slightly lower for larger window sizes, as depicted in Figure 2c. *FS-4* performs worse than the other sensors. For combinations of 8 sensors, as

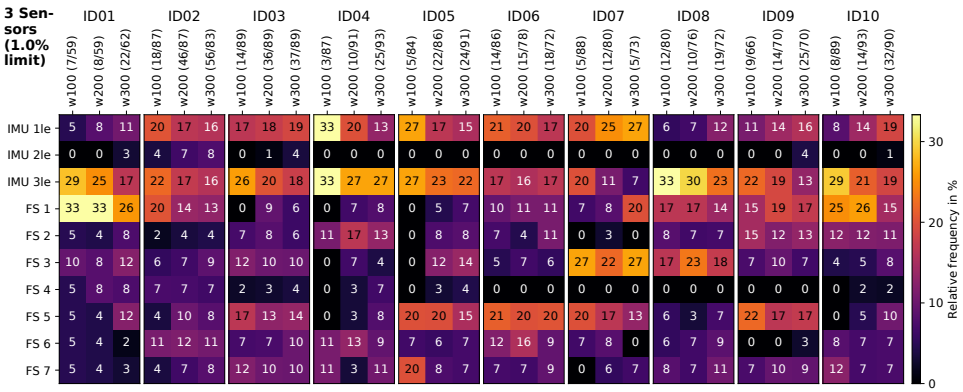


Fig. 3: Heatmap for the single subject evaluations with different window sizes and 3 sensors.

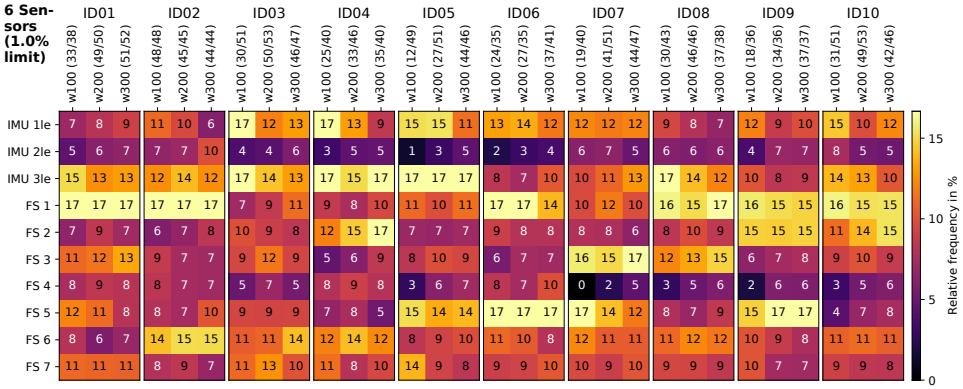


Fig. 4: Heatmap for the single subject evaluations with different window sizes and 6 sensors.

shown in Figure 2d, the differences of occurrences between the sensors get smaller, since only a few sensors can be left out. *FS-4* performs still worse than the other sensors.

For the evaluation of the multi-subject data it can be seen in total that IMUs occur more frequently than force sensors. For combinations of up to 6 sensors, *IMU-1le* and *IMU-3le* perform better than *IMU-2le*, thereafter the frequency of occurrence of sensor *IMU-2le* is equal. The frequency of occurrence of *FS-4* is considerably lower than for the other force sensors. Apart from this exception, the differences among the force sensors are relatively small and vary with the window size. The worse performance of *FS-4* might be explained by some measurements being unrealistically high.

With all 10 sensors and a window size of $300\,\mathrm{ms}$ we achieve an accuracy of $92.40\,\%$ in [8]. With our exploration of the feature space approach we achieve an F_1 score of $90.95\,\%$. We attribute the slight difference in this paper to some parameters being adapted, such as the number of stratified cross fold validations and the number of training steps. The adaptations lead to a reduced training time.

*2) **Single-Subject**:* An excerpt of the evaluation of the single-subjects is depicted in Figure 3 and Figure 4.

For each single subject a heatmap of the frequency of occurrences for each sensor modality was generated. By using 3 features combined (Figure 3) it can be seen that *IMU-2le* is mostly not used, whereas *IMU-1le* and *IMU-3le* appear very often. Regarding the force sensors, some force sensors, e.g. *FS-1*, play an important role for subject $1, 2, 8, 9$ and 10 while for the other subjects other force sensors are more important, e.g. *FS-3* and *FS-5*. Force sensor *FS-4* performs worst for most subjects.

Using 6 sensors combined a usage of all sensors can be seen (Figure 4). *IMU-2le* and *FS-4* perform still worse compared to the other sensors but become more relevant. At this point *FS-2* and *FS-6* are more used than before. Depending on the subject, different sensors remain most common.

In summary, it can be said (not depicted by graphics) that for groups of up to two sensors, the high suitability of *IMU-1le* and *IMU-3le* stands out. From 3 sensors upwards depending on the subject individual force sensors are well suited, in particular *FS-1, FS-3* and *FS-5*. *IMU-2le* and *FS-4* remain particularly rare. *FS-2* and *FS-6* frequently occur at a later stage. The difference between the more frequently and less frequently occurring sensors decreases with increasing dimension. Using 9 sensors there are only slight differences between the sensors. The dominance of sensors can be seen as a continuous development across the different sensor numbers, in contrast to a distribution being completely different for each sensor number. There are some differences in the frequency of a sensor for a certain subject between different window sizes.

*3) **Classification Performance per Motion Type**:* A further important aspect is the classification performance per

motion type. For future analyses and improvement of future recordings it is important to know which movements cause a worse classification performance. For that purpose, we evaluated the data of subject 1 in more detail.

Figure 5 shows the aggregated results of the features *IMU-1le* and *IMU-3le*, which are the suitable feature sets according to the evaluation process used for the heatmaps. The confusion matrix presents the data of subject 1 and a window size of 100 ms. The rows correspond to the performed motion, whereas the columns represent the classified motion. The shortcuts of the motions are in line with those introduced in Section III-B.

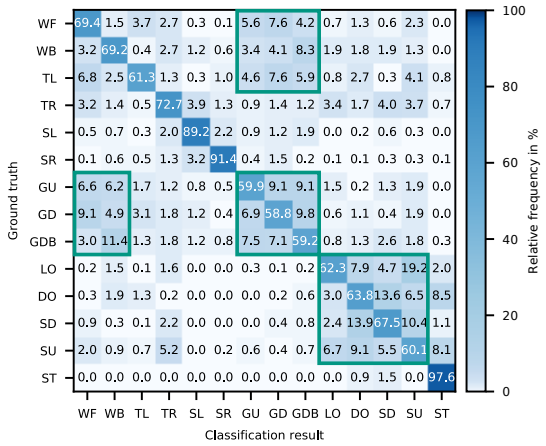


Fig. 5: Aggregated confusion matrix of subject 1.

The motion classes which are often mixed up are highlighted in Figure 5. *Walking Forward (WF)* and *Walking Backward (WB)* are often misclassified as *Going Upstairs (GU), Going Downstairs (GD)* and *Going Downstairs Backwards (GDB)* or otherwise. Additionally, *Lift Object (LO)* and *Drop Object (DO)* are often mixed up with *Stand Up (SU)* and *Sit Down (SD)*. The results are consistent since the motion types are very similar in the execution.

Analysing only single sensors, a large difference between force and IMU sensors can be observed. This is shown in Figure 6 and Figure 7.

IMU-2le achieves an overall F_1 score of 58.13 %, *FS-1* achieves 60.33 %. Since the scores are nearby, these sensors were chosen for the comparison. *IMU-2le* performs worst among the IMU sensors and *FS-1* performs best among all force sensors. Figure 6 shows that only the motion *Stand (ST)* is reliably classified. Especially clear are the differences of the motion *Lift Object (LO)* which achieves very low results for *IMU-2le* namely 9.5 %, while *FS-1* enables correct detection in 79.1 % of the cases. It should be noted at this point that the other two IMU sensors achieve significantly better results than *IMU-2le*, even though they remain below the values of most force sensors. Other motions such as *Sidestep Right (SR)* or *Sidestep Left (SL)* achieve better results when using *IMU-2le* instead of *FS-1*. Depending on the modality, different movements are easily distinguishable. However, even within the same modality, the differences for some classes are high.

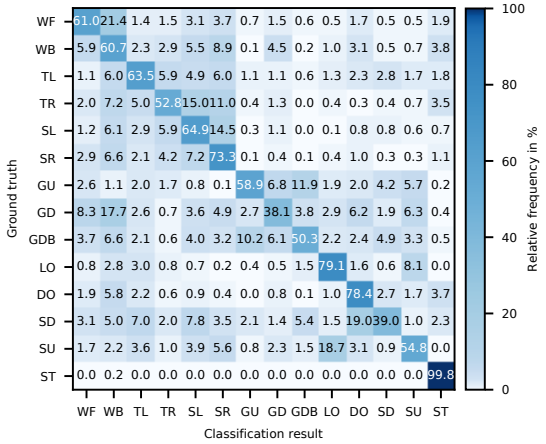


Fig. 6: Confusion matrix of subject 1 using only FS-1.

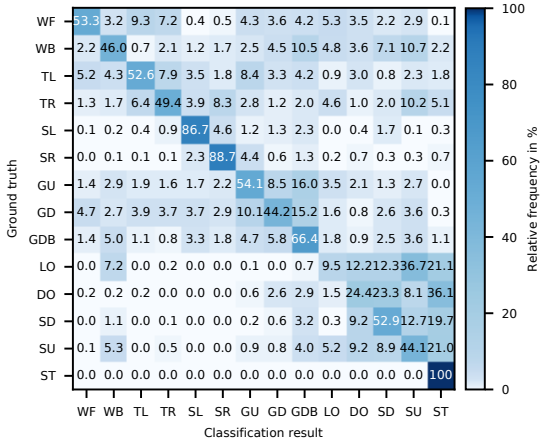


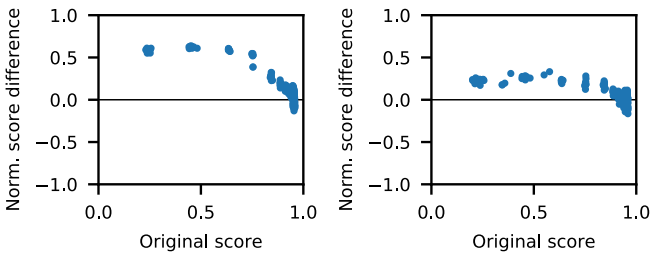
Fig. 7: Confusion matrix of subject 1 using only IMU-2le.

B. Fine-grained Reduction

An analysis using more fine-grained features was conducted for two single subjects using the window size 100 ms. In contrast to the sensor-based analyses, the features are split into smaller parts. The force sensor data is split in its x, y and z direction to verify if some directions play a more important role or if even a one-dimensional force sensor would be sufficient. The feature equals therefore $f_1(B)$. The y direction of the force sensor directs upwards, the z direction towards the leg and the x direction respectively to build a right-handed coordinate system. Technically, splitting the Euler angles and linear acceleration values of the IMU sensors in their single directions would be possible as well. However, as Euler angles are sometimes ambiguous, providing just a single scalar value is meaningless. Therefore, the features are defined as $f_2(A)$.

For the evaluation, the following procedure was used: Adding one feature to a feature set results in a difference of the F_1 score. To account for the overall score being limited by 100 %, this difference is normalized by the highest achievable difference. Figure 8 depicts the impact of *IMU-1e* and *FS-1z* for subject 1 and a window size of 100 ms. While the normalized difference being approximately constant for

lower original scores, it drops for very high ones until it scatters around zero. The drop might be caused by information already being provided by the original feature set (redundancy), leading to fewer added information and smaller improvements. The scattering might be due to training being a randomized process, which gets dominant as soon as the achievable differences become very small. When interpreting such plots, one has to assure a lack of values scattering around zero is not caused by the absence of data points with high original scores, which usually occur when the original feature set is of a high dimension. When comparing normalized score differences for the same original score, it can be seen that *IMU-1le* achieves higher improvements than *FS-1z*, but in both cases, the improvement is limited at a certain original score.



(a) IMU-1e (b) FS-1z

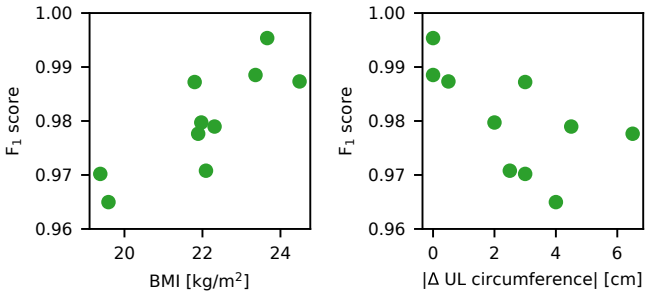
Fig. 8: Normalized score difference of ID 1 with window size $100\,\mathrm{ms}$.

By comparing the single features against each other (not depicted here in detail), it can be seen that for the y direction *FS-1y* and *FS-2y* perform best, followed by *FS-4y, FS-6y* and *FS-7y*. For the x direction *FS-2x* and *FS-5x* reveal good results, thereafter *FS-1x, FS-4x* and *FS-6x*. Regarding the z direction *FS-1z, FS-2z, FS-3z, FS-4z* and *FS-6z* perform equally good. Summarized, it can be observed that *FS-1* and *FS-2* perform best, *FS-4* and *FS-6* are thereafter. The worst results are achieved with *FS-3, FS-5* and *FS-7*. Compared to the frequency of occurrence results of Section V-A.2 the results of both analyses are similar for *FS-1* and *FS-2*, while the results differ especially for *FS-3, FS-5* and *FS-7*. Here, it can be seen that the single directions play an important role for the classification performance. Within one sensor modality, different directions perform best and this also differs between subjects. Therefore, it is difficult to favour certain directions. This leads us to the assumption that it is worthy to use 3D force sensors in context of motion classification with exoskeletons instead of 1D sensors.

Among the IMU sensors, *IMU-3e* achieves better results that the other features. In general, the Euler angles perform better than the linear accelerations. Nevertheless, the linear accelerations contribute to a notable improvement of the classification results and should be utilized as well.

C. Influence of Physical Characteristics

Since the classification performance in Section V-A differs between the subjects, we analysed the influence of their physical characteristics. Further details of our subject group are listed in [7] Section IV-A. Figure 9 depicts the F_1 scores for individual subjects depending on their Body Mass Index (BMI) and the difference of the upper leg (UL) circumference to the exoskeleton's nominal value. The results were retrieved using all sensors and a window size of $300\,\mathrm{ms}$. The average BMI of the subjects was about $22\,\mathrm{kg/m^2}$ (Std. dev. 1.5) and the average UL circumference $55.2\,\mathrm{cm}$ (Std. dev. 3.2).



(a) BMI values of single subjects. (b) UL differences within subjects.

Fig. 9: Classification performance of all subjects depending on their physical characteristics.

Figure 9a shows that the classification performance increases with the BMI. This might be caused by the rigid parts of the exoskeleton. The exoskeleton can be tightened to the user only to a certain extent. If a subject is too thin it could happen that the force sensors do not have enough contact to measure the interaction forces correctly.

Figure 9b depicts the deviation of a subject's upper leg circumference to subject 2, whom the exoskeleton was designed for. A larger deviation correlates with a decrease of the classification accuracy. This indicates that a good fit and correct tightening of the exoskeleton to the user play a crucial role for the motion classification when using force sensors. In future exoskeletons it should be possible to better adjust the exoskeleton to the user.

VI. Conclusion

In this paper, we investigated further whether the number of sensors or features used to classify motions of our unilateral, passive lower limb exoskeleton can be reduced while maintaining the classification performance. For this purpose, we used the data of 10 subjects who performed 14 different daily activities wearing our exoskeleton which is equipped with three IMUs and seven 3D force sensors. We evaluated the data with our motion classification approach based on Hidden Markov Models (HMMs) as introduced in [7].

We presented our adaptation of our systematic exploration of the feature space approach ([9]) in combination with our HMM-based motion classification approach. Applying this exploration approach reduces the risk of discarding important feature sets and decreases the training and testing time. Our evaluation addressed first the problem of a sensor-based dimensionality reduction. We showed that when training and testing with multi-subjects the IMU sensors performed better than the force sensors but using both sensor modalities in

combination increases the classification performance significantly. *IMU-2le* which is located on the lower leg performs worse than the other IMU sensors. *FS-4* achieves worse results than the other force sensors. We can reduce our sensor setup to 7 sensors, e.g. to *IMU-1le, IMU-2le, IMU-3le, FS-1, FS-5, FS-6* and *FS-7* while still maintaining about the same classification accuracy. In the context of the sensor-based dimensionality reduction we also performed an evaluation with training and testing on single-subjects. The results show that *IMU-1le* and *IMU-3le* sensors achieve better results than the *IMU-2le*. Depending on the subject, different force sensors revealed good results. Regarding the classification accuracy of single motion types we observed that for some motions, e.g. *Sidesteps Left*, the IMU sensors achieve better classification results and that for other motions, e.g. *Lift Object*, the force sensors perform better.

In addition, we conducted a more fine-grained dimensionality reduction where the features were split into smaller parts. The results show that different force directions or IMU sensor values were preferred depending on the subject. Therefore, we suggest to use both, the linear accelerations and Euler angles of the IMU sensors and three dimensional force sensors. Furthermore, we looked at the subjects' physical characteristics to determine their influence on the motion classification. A higher BMI and a smaller deviation of the upper leg circumference lead to higher classification results.

In our future work we will address the question of using derived features to determine if the classification performance can be increased and the training time can be minimized. Furthermore, we will investigate how the motion classification framework can be used to support the prediction of the next possible motion type in a given context to allow a faster adaptation to the user's movement.

REFERENCES

[1] S. Herath, M. Harandi, and F. Porikli, "Going deeper into action recognition: A survey," *Image and vision computing*, vol. 60, pp. 4–21, 2017.

[2] G. Cheng, Y. Wan, A. N. Saudagar, K. Namuduri, and B. P. Buckles, "Advances in human action recognition: A survey," *arXiv preprint arXiv:1501.05964*, 2015.

[3] A. Jordao, A. C. Nazare Jr, J. Sena, and W. R. Schwartz, "Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art," *arXiv preprint arXiv:1806.05226*, 2018.

[4] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE communications surveys & tutorials*, vol. 15, no. 3, pp. 1192–1209, 2012.

[5] C. Chen, R. Jafari, and N. Kehtarnavaz, "A survey of depth and inertial sensor fusion for human action recognition," *Multimedia Tools and Applications*, vol. 76, no. 3, pp. 4405–4425, 2017.

[6] J. Beil, G. Perner, and T. Asfour, "Design and control of a lower limb exoskeleton (kit-exo-1)," in *Int. Conf. on Rehabilitation Robotics (ICORR)*, pp. 119–124, 2015.

[7] J. Beil, I. Ehrenberger, C. Scherer, C. Mandery, and T. Asfour, "Human motion classification based on multi-modal sensor data for lower limb exoskeletons," in *2018 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 5431–5436, IEEE, 2018.

[8] I. Patzer and T. Asfour, "Minimal sensor setup in lower limb exoskeletons for motion classification based on multi-modal sensor data," in *Int. Conf. on Intelligent Robots and Systems (IROS)*, IEEE, 2019.

[9] C. Mandery, M. Plappert, J. Borras, and T. Asfour, "Dimensionality reduction for whole-body human motion recognition," in *19th Int. Conf. on Information Fusion (FUSION)*, pp. 355–362, IEEE, 2016.

[10] A. Malaisé, P. Maurice, F. Colas, and S. Ivaldi, "Activity recognition for ergonomics assessment of industrial tasks with automatic feature selection," *IEEE Robotics and Automation Letters*, 2019.

[11] S. Feldhorst and M. ten Hompel, "Bewegungsklassifikation mithilfe mobiler Sensoren zur Analyse des Kommissionierprozesses," *Logistics Journal: Proceedings*, 2016.

[12] J. Jang, K. Kim, J. Lee, B. Lim, J.-K. Cho, and Y. Shim, "Preliminary study of online gait recognizer for lower limb exoskeletons," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 5818–5824, IEEE, 2017.

[13] A.-C. Tsai, J.-J. Luh, and T.-T. Lin, "A novel stft-ranking feature of multi-channel emg for motion pattern recognition," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3327–3341, 2015.

[14] J. Gancet, M. Ilzkovitz, E. Motard, Y. Nevatia, P. Letier, D. de Weerdt, G. Cheron, T. Hoellinger, K. Seetharaman, and M. Petieau, "Mindwalker: going one step further with assistive lower limbs exoskeleton for sci condition subjects," in *IEEE RAS & EMBS Int. Conf. on Biomed. Rob. and Biomechatronics (BioRob)*, pp. 1794–1800, 2012.

[15] A. Kilicarslan, S. Prasad, R. G. Grossman, and J. L. Contreras-Vidal, "High accuracy decoding of user intentions using eeg to control a lower-body exoskeleton," in *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5606–5609, 2013.

[16] A. H. Do, P. T. Wang, C. E. King, S. N. Chun, and Z. Nenadic, "Brain-computer interface controlled robotic gait orthosis," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, p. 111, 2013.

[17] A. Villa-Parra, D. Delisle-Rodríguez, A. López-Delis, T. Bastos-Filho, R. Sagaró, and A. Frizera-Neto, "Towards a robotic knee exoskeleton control based on human motion intention through eeg and semgsignals," *Procedia Manufacturing*, vol. 3, pp. 1379–1386, 2015.

[18] T. Lorrain, N. Jiang, and D. Farina, "Influence of the training set on the accuracy of surface emg classification in dynamic contractions for the control of multifunction prostheses," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, no. 1, p. 25, 2011.

[19] J. Liu, D. Zhang, X. Sheng, and X. Zhu, "Quantification and solutions of arm movements effect on semg pattern recognition," *Biomedical Signal Processing and Control*, vol. 13, pp. 189–197, 2014.

[20] D. Yang, Y. Gu, L. Jiang, L. Osborn, and H. Liu, "Dynamic training protocol improves the robustness of pr-based myoelectric control," *Biomedical Signal Processing and Control*, vol. 31, pp. 249–256, 2017.

[21] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.

[22] A. Y. Yang, S. Iyengar, S. Sastry, R. Bajcsy, P. Kuryloski, and R. Jafari, "Distributed segmentation and classification of human actions using a wearable motion sensor network," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 1–8, 2008.

[23] O. D. Lara, M. A. Labrador, *et al.*, "A survey on human activity recognition using wearable sensors.," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[24] A. Malaisé, P. Maurice, F. Colas, F. Charpillet, and S. Ivaldi, "Activity recognition with multiple wearable sensors for industrial applications," in *Advances in Computer-Human Interactions*, 2018.

[25] L. Steffan, L. Kaul, and T. Asfour, "Online stability estimation based on inertial sensor data for human and humanoid fall prevention," in *IEEE-RAS 17th Int. Conf. Humanoid Robotics*, pp. 171–177, 2017.

[26] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas, "Window size impact in human activity recognition," *Sensors*, vol. 14, no. 4, pp. 6474–6499, 2014.

[27] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 1, pp. 111–118, 2007.

[28] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 2, pp. 186–192, 2010.

[29] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics," *Robot. Auton. Syst.*, vol. 73, pp. 155–170, 2015.