Minimal Sensor Setup in Lower Limb Exoskeletons for Motion Classification based on Multi-Modal Sensor Data

Isabel Patzer and Tamim Asfour

***Abstract*— Exoskeletons are considered to be a promising technology for assisting and augmenting human performance. A number of challenges related to design, intuitive control and interfaces to the human body must be addressed. In this paper, we approach the question of a minimal sensor setup for the realization of control strategies which take into account the actions currently performed by the user. To this end, we extend our previous work on online classifications of a human wearing a lower limb exoskeleton in two directions. First, we investigate the minimal number of sensors that should be attached to the exoskeleton to achieve a certain classification accuracy by investigating different sensor setups. We compare results of motion classification of 14 different daily activities such as walking forward and going upstairs using Hidden Markov Models. Second, we analyse the influence of different window sizes, as well as the classification performance of different motion types when training on multi- and single-subjects. Our results reveal that we can reduce our sensor setup significantly while achieving about the same classification performance.**

I. Introduction

The area of augmenting exoskeletons moved into focus in research as well as in industrial applications in recent years. Such devices can be used for example to enhance human's motor abilities. Apart from their mechatronics design, the interface to the human body and the control strategies for exoskeletons are the key for intuitive and effortless operation by the user. In this work, we address the following research question: what is the minimal sensor setup for a lower limb exoskeleton which is needed for a reliable recognition and classification of human's actions?

In our previous work, we presented a lower limb exoskeleton (KIT-EXO-1) with two active Degrees of Freedom (DoF) with a concept for a force-based interface to the human body [1]. In [2], we introduced a Hidden Markov Model (HMM) based motion classification system and evaluated it using an unilateral, passive lower limb exoskeleton for the left leg, which is equipped with seven 3D force sensors arranged in a way to correspond to the main muscles in the human leg to allow robust capturing of interaction forces between the exoskeleton and the human lower limb. Furthermore, 3 Inertial Measurement Units (IMUs) are attached on the device, one on each segment. We evaluated the online classification performance with multi-subject data using a sliding-window approach. We further investigated the latency and generalization performance of our approach with multi-subject data and different window sizes.

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

The authors are with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Germany.
{isabel.patzer, asfour}@kit.edu

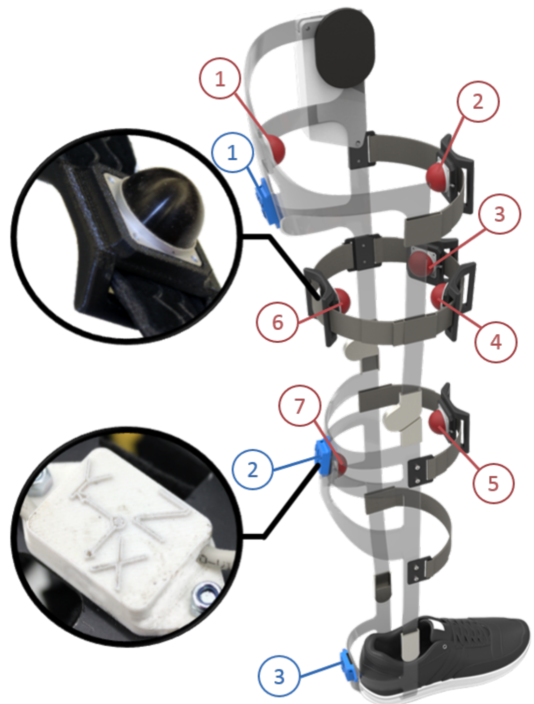

Fig. 1: Passive exoskeleton with 7 3D-force (red) and 3 IMU sensors (blue). The numbers correspond to the sensor labels.

In this paper, we present an evaluation comparing multi- and single-subject classification performance using different window sizes. Furthermore, we analyse the classification accuracies of single motion types and address the question of a minimal set of sensors required to achieve a certain classification accuracy.

The paper is organized as follows. Section II covers a brief overview of sensor setups and machine learning methods and their combinations used in the area with and without exoskeletons as well as different feature reduction approaches. In Section III, we give a short recap of our previous work and an overview of different feature selection methods. Section IV covers our evaluations and results, while Section V offers the conclusion and outlook.

II. Related Work

The problem of finding an optimal sensor setup in combination with a suitable human motion recognition and classification system was addressed by several authors in the literature. Tsai et al. [3] studied upper arm motion patterns using multi-channel EMG signals and Support Vector

Machines (SVMs) for the control of exoskeleton robots. EEG-based human-robot interfaces also find application for controlling different lower limb exoskeletons ([4], [5], [6]). The H2 exoskeleton is controlled using identified patterns from EEG and sEMG signals, based on Artificial Neuronal Network and SVM classification [7].

One disadvantage of using EEG or EMG sensors is the loss of classification accuracy due to the individualized signal patterns of each subject ([8], [9], [10]). Therefore, mechanical sensors, such as Inertial Measurement Units (IMUs) or torque sensors, are often used for motion classification, with and without exoskeletons. It is especially common in industrial environments to place inertial sensors on the human body ([11], [12], [13]).

Attal et al. [11] compared 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. used a wearable motion tracking suit (MVN Link suit Xsens) consisting of 17 IMUs and a sensorized glove (Emphasis Telematics) that processed force data for motion classification via HMMs [14], [15]. Their work focuses on industrial applications.

Taborri et al. [16] used force and IMU sensors to detect the gait phase of a lower limb orthosis. Wang et al. [17] classified motions with SVMs and IMU data and predicted possible upcoming motions with HMMs for the Non-Binding Lower Extremity Exoskeleton (NBLEX). Gong et al. [18] proposed a real-time, on-board training and recognition method for an active pelvis orthosis using two IMUs.

How to reduce the number of sensors and features in order to scale down computational effort while keeping high classification and prediction accuracy is a common problem. A widely used strategy is to apply feature selection algorithms that retrieve only the relevant features for the motion classification process. In our previous work [19], we used a wrapper-based method for whole-body human motion recognition based on kinematic data using HMMs for motion classification and showed that a lower-dimensional feature spaces is sufficient to achieve high motion recognition performance (4 dimensions and 97.76% accuracy).

Malaisé et al. [14] made comparisons amongst Principal Component Analysis (PCA), feature-based, and wrapper-based methods for their HMM-based motion classification approach using data from a wearable motion tracking suit and a sensorized glove. In their analysis of data from 13 subjects, they showed that a wrapper-based method performed best compared to the other two.

III. Sensor System and Motion Classification

We provide a brief overview of our previous works ([2], [19]) on motion classification, system setup, data and methods which are used for the evaluations in Section IV.

A. Exoskeleton and Sensors

Our passive lower limb exoskeleton for the left leg (see [2] and Figure 1) was also used for this work. The frames of the exoskeleton cover the thigh, shank and foot. The connection between these three components consists of orthotic revolute joints[1]. To wear the exoskeleton, orthotic Velcro straps are used on the anterior thigh and shank.

Orientations and linear accelerations of the thigh, shank and foot are measured with a total of 3 IMUs[2], one on each limb segment. The data is recorded at a frequency of 80 Hz. To measure the interaction forces between the exoskeleton and the wearer 7 individual 3D force sensors[3] are placed over the large muscles of the front and back of the thigh, as well as on the shank. The raw data of the force sensors are collected with a maximum frequency of 100 Hz.

B. Data

Our previous collected data described in [2] consists of 13 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)* and *Sit Down (SD)*. Based on our results, we also decided to add the motion *Stand (ST)* to our motion classes. This motion has already been recorded in our previous study but was not used. This left us with a total of 14 different motion types. Every motion was recorded 10 times, each performed by 10 subjects (5 male, 5 female). Since the IMU data (80 Hz) and the force data (100 Hz) were recorded at a different frequency, the IMU values were interpolated to 100 Hz and the timestamps were unified.

C. Motion Classification

The evaluations and results reported in Section IV are based on a Hidden Markov Model (HMM) multi-class classification. We trained one HMM for each motion type, leading to 14 HMMs in total. Our HMMs were trained with a fully connected topology and diagonal covariance matrices. The number of states for each HMM added up to 14 states and Gaussian distributions were used to model observations. To ensure online application, we use a sliding window approach - splitting training and test data into windows of 100, 200, 300, 400 or 500 ms (depending on our evaluation). We start a new window every 10 ms. A currently tested window is assigned to the HMM with the highest log-likelihood.

As training input, we used the values of the force and IMU sensors. 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 a dimensionality for the IMU feature vector of 18 (containing all 3 IMUs). For further information about our motion classification approach, we refer to [2].

[1] OttoBock HealthCare; 17B47=20 / 17B57=20

[2] BNO055 IMU, Robert Bosch GmbH

[3] Optoforce Ltd. OMD-30-SE-100N

D. Feature Selection

There are different methods for retrieving the optimal feature set for the chosen motion classification system. One approach is to test every possible feature or sensor combination (*brute-force* approach). Depending on the number of features or sensors, this is not always feasible, due to long training times. A solution for this problem are feature selection methods. These algorithms are used to reduce the dimensionality of the input data for motion classification. A subset of a high dimensional feature set is selected realizing the best classification performance. There are three types of methods which are often used in the context of machine learning: embedded, filter-based, and wrapper-based methods [20].

Embedded methods select the features based on the results of the training process and are therefore specific to a chosen machine learning algorithms. Filter-based methods assign a score to the features individually, regarding the observation, based on a given filter measure. There is a large number of such filter methods available [21]. Filter-based methods are not classifier-dependent and therefore not time-consuming since no training of the classifier is needed. A drawback is that the algorithm does not know relationships amongst the features, which could lead to selection of redundant features. Wrapper-based methods [22] are based on the performance of a classifier within the tested features. In contrast to embedded methods, this approach evaluates the classification performance and not the classifier itself. Here, a strategy for searching through the feature space must be chosen. In contrast to filter-methods, this method can consider relationships between features. Since the chosen feature subsets are based on classification results, this method has a higher computational effort as embedded or filter-based methods.

In our approach we use the wrapper-based method we developed in our previous work [19], where we conducted a feature space dimensionality reduction for the recognition of whole-body human actions based on Hidden Markov Models. Here, the total number of features is denoted as N. The chosen classifier (in this case HMM) is trained independently for each single feature combination. Afterwards, the outcome of the trained model is evaluated with a chosen metric which has to be set before. The best features of amount M are retrained in combination with one other feature. For every iteration, one feature is added to this combination until the chosen maximum dimensionality of feature combinations is reached. Additionally, several feature subsets are kept at each iteration to reduce the risk of eliminating important subsets. For the evaluation of the given feature subset, a stratified 3-fold cross validation is used. For further information, we refer to [19].

Due to our sensor setup the total number of features N is denoted as 10 since our features corresponds to the sensor values introduced in Section III-B. As chosen in [19] we also set the best amount of features $M = 10$. To reduce computational costs we scaled our stratified 5-fold cross validation down to a stratified 3-fold cross validation as in [19]. Furthermore, we reduced the amount of trainings step as proposed in [19].

IV. Evaluations

We first give a more detailed overview regarding the results of our single subject evaluations compared to our previous presented results in [2] before introducing the motion classification performance per motion type. Then, we investigate if there are specific motion combinations often mixed up in the classification. Finally, we address the question of a minimalistic sensor setup.

A. Single Subject Evaluation

To better compare ourselves to other motion classification approaches, we performed further analysis regarding single subject evaluations with window sizes of 100-500 ms since many other approaches are just training and testing on single-subjects and not multi-subjects. We performed a stratified 5-fold cross validation using the data of all 10 sensors of our passive exoskeleton for our HMM-based motion classification as presented in Section III-C. In this evaluation we only used the data of 13 motion types (see Section III-B) to directly compare our new results to the ones of our previous work [2].

The results can be seen in Table I. The first column depicts the window size, the second column the average classification accuracy when training on all subjects (AS) and the last column the average classification accuracy when training on single-subjects (SS).

TABLE I: Accuracies for different window sizes comparing single- and multi-subject evaluations.

Window Size [ms]	Accuracy AS [∅%]	Accuracy SS [∅%]
100	82.92	95.35
200	89.16	97.39
300	92.80	98.36
400	95.00	98.95
500	96.46	99.33

When all 10 subjects (AS) are contained in the training set, the average accuracy of window size 100 ms adds up to 82.92 %. With an increasing window size, the accuracy also increases since the observing time of the windows also raises and thereby the data gets less unambiguous (see also [2]). By running a single-subject (SS) evaluation where only one subject is considered in the training and testing set for every evaluation, the classification accuracy is already very high at the beginning and increases slower for bigger window sizes.

Comparing both evaluations shows that the average accuracy of the SS evaluations increase enormously for the window sizes 100, 200 and 300 ms compared to the AS evaluation. This reveals that our approach of using a combination of force and IMU sensors can achieve a high accuracy of 95.35 %, even for a window size of 100 ms. Due to the small window size the approach is therefore online applicable.

B. Classification Performance per Motion Type

In our next evaluation we wanted to analyse in detail how the classification performance of each single motion is conducted for training on all subjects combined or just training on a single subject. This evaluation was performed with all 14 motion classes, a stratified 5-fold cross validation and a window size of 100-500 ms (leaning on our results of [2]).

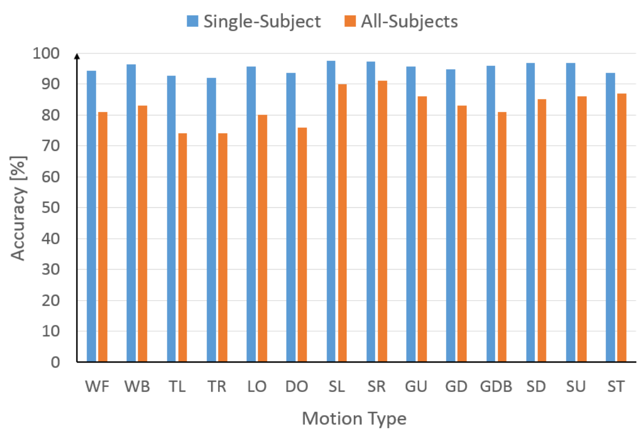

Fig. 2: Classification accuracy per motion type using a window size of 100 ms.

Figure 2 shows the results for the analysis of window size 100 ms. Blue depicts the single-subject evaluations (SS) and orange the evaluations over all subjects combined (AS). Every bar group corresponds to one motion type. The abbreviations for every motion type are listed in Section III-B. The height of every bar corresponds to the accuracy of the motion classification. Figure 2 reveals that the classification accuracy of the SS evaluation (blue) achieve high results for all motion types. Between the different motion types there is only a slight difference, only *Turn Left (TL)* and *Turn Right (TR)* have a slightly lower classification accuracy as the other motions.

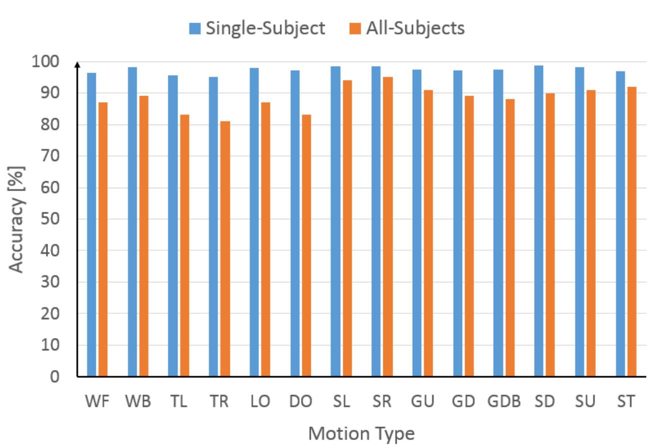

Fig. 3: Classification accuracy per motion type using a window size of 200 ms.

The classification accuracy of the AS evaluation (orange) in Figure 2 is much less (as also stated in Section IV-A). Here, a widely scattered performance between the single motion types can be observed. The motions *Turn Left (TL), Turn Right (TR)* and *Drop Object (DO)* have the lowest classification accuracies around 75 %. *Sidesteps Left (SL)* and *Sidesteps Right (SR)* have the highest accuracy with about 90 %.

Figure 3 depicts the results when running the same evaluation for a window size of 200 ms. The SS evaluation can achieve overall higher accuracies as with using a window size of 100 ms. The motions *Turn Left* and *Turn Right* perform again slightly worse. The classification accuracies in the AS evaluation retrieved as well better results. The difference between *Going Downstairs (GD), Going Upstairs (GU)* and *Going Downstairs Backwards (GDB)* become less. The same behaviour can be observed for the window sizes 300, 400 and 500 ms and are therefore not depicted here in detail.

C. Mixed Up Motions

Based on the results of Section IV-B we wanted to further investigate which motion was classified and what real motion was actually performed (mixed up motion). We were interested if there are certain motion combinations causing a worse classification performance or if there is an overall tendency of wrong classified motion pairs. This evaluation was also conducted with all 10 subjects, all 14 motion classes, a stratified 5-fold cross validation and window sizes of 100-500 ms. Every presented evaluation is an excerpt of all mixed up motion combinations. For every window size the top 5 combinations are shown.

TABLE II: Comparison of real performed and classified motions for window size 100-500 ms. Abbreviations: see Section III-B

WS [ms]	Classified Motion	Real Motion	Amount [%]
100	WB	GDB	4.81
	GD	WF	4.42
	TL	WF	3.86
	GDB	WB	3.41
	DO	SD	3.13
200	WB	GDB	6.85
	GD	WF	5.05
	TR	WF	3.80
	TL	WF	3.63
	DO	SD	3.32
300	WB	GDB	10.12
	GD	WF	5.68
	TR	WF	4.95
	TL	WF	4.36
	DO	SD	3.81
400	WB	GDB	12.84
	GD	WF	7.21
	TR	WF	5.69
	TL	WF	4.82
	DO	SD	3.97
500	WB	GDB	14.77
	TR	WF	8.73
	GD	WF	8.67
	TL	WF	5.26
	DO	SD	4.62

Table II depicts the results. The first columns corresponds to the analysed window size, the second column contains the classified motion, the next one the real performed motion and the last column depicts how often this combination appears amongst all wrong classified motions. The classification accuracy for every window size is shown in Table I in Section IV-A.

For every window size almost the same observations can be made. The most mixed up motion combinations for every window size are *Walking Backward (WB)* and *Going Downstairs Backward (GDB)*. The underlying motion is nearly the same despite the difference that the *Going Downstairs Backward* motion has a larger vertical component. This is similar to *Walking Forward* and *Going Downstairs* which as well is mixed up. Here, the vertical component of the motion should be taken into account to improve the classification accuracy. One solution to this could be to rate the vertical component stronger.

The confusion of *Turn Right/Left (TR/TL)* and *Walking Forward (WF)* could arise due to the fact that the motions *Turn Right/Left* have locomotion elements of *Walking Forward* since the subject was asked to walk a $90°$ motion.

The motion types *Drop Object (DO)* and *Sit Down (SD)* have also similar locomotion elements since during sitting down a squat motion is performed which was also conducted when dropping the box (back-friendly dropping of the object and no leaning forward dropping).

For smaller window sizes the amount of wrong classified pairs is in general smaller and more widely dispersed as with bigger window sizes. For window size $300\,\mathrm{ms}$ and larger the differences became more significant. One solution could be to cut the recordings to remove ambiguous parts such as in the motions *Turn Right/Left*. Nevertheless, it is more relevant to apply more intelligent solutions to these problems since there are always elements of other motions integrated. A further approach could be to implement a high level classification strategy which takes for example the last n classifications into account or reduces the calculated classification accuracy for unlikely motion transitions, e.g. if the person is sitting it is unlikely that the person will perform any kind of walking motion next.

D. Reduced Sensor Setup

Our final analysis addresses the question of a minimalistic sensor setup resulting from reducing the number of sensors in our sensor setup while still achieving a high classification accuracy. In Figure 1 the labels of the sensors are depicted. The force sensors are highlighted in red and the IMU sensors are coloured blue. To this end, we first investigate a reduction of number of sensors using a *brute-force* approach. In our case we have 10 sensors from which we can choose.

This sensor setup is composed of seven 3D force sensors which deliver force information in x,y and z direction and three IMU sensors, each of which provides linear acceleration and RPY-angle information. With 10 possible sensors, we retain 1023 possible combinations to test. These evaluations were conducted with all 10 subjects, all 14 motion classes, a stratified 5-fold cross validation and a window size of $300\,\mathrm{ms}$ (based on our results of [2]). With this window size we can achieve with all 10 sensors combined a classification accuracy of $92.80\,\%$ (see Table I, using 13 motions). We conducted the *brute-force* analysis for all possible combinations and we discuss here important aspects of these results.

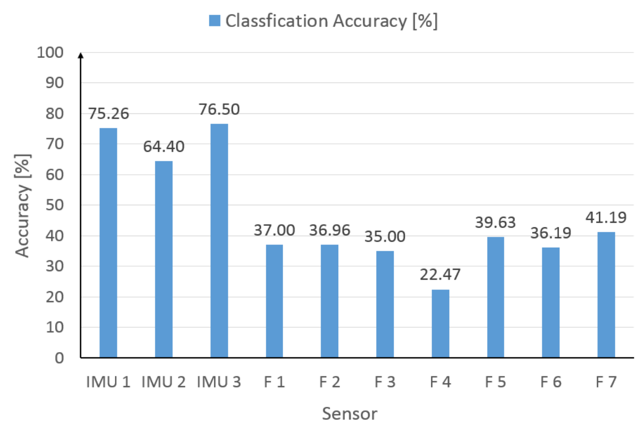

Fig. 4: Classification accuracies using single IMU or single force F sensors.

Figure 4 reveals that by using just one single IMU sensor we achieve a classification accuracy in the range of $64.40\,\%$-$76.50\,\%$. The best result with IMU sensor 3 is just around $16\,\%$ less than the result of using all 10 sensors combined. The IMU sensors 1 and 3 perform similar, however IMU sensor 2 performs much worse as a single used sensor. The force sensors, denoted as F, can just achieve an accuracy in the range of $22.47\,\%$-$41.19\,\%$.

Table III shows the classification performance of different sensor combinations. For each amount of sensor with different sensor combinations, the top 10 sensor combinations are listed. The first column shows the amount of used sensors, the second and third column which force and IMU sensors are used and the last column the classification performance. Using two sensors combined the classification accuracy increases. Mostly, one force sensor and one IMU sensor are combined. Regarding the results of the single sensor evaluation, the assumption could be that combining two IMU sensors would achieve the highest classification accuracies. This is the case for a combination of IMU sensor $1, 3$ and $1, 2$ but the combination of IMU sensor $2, 3$ can just achieve $79.49\,\%$ (not shown in Table III since this combination does not belong to the top 10 sensor combinations). Starting from dimension 2, the force sensors move more into focus.

In combination with force sensors, only IMU sensor 1 and 3 appear. The force sensors 5 and 7 are used more often than the others but there is no such strict tendency of upcoming sensors such as with the IMU sensors. The difference of the classification accuracy between the top 5 single sensor pairs are very small (between $0.01\,\%$-$2.06\,\%$). This occurs for all sensor pairs from 2 up to 9 sensors combined. Therefore, an absolute assertion about the best sensor combination per dimension can not be made, but rather a statement about tendencies which sensors often appear.

Also with adding more sensors the best results are achieved with a combination of two or more IMU sensors. The combination of IMU sensor 1 which is located on the upper leg, with IMU sensor 3 which is located on the foot provides the best results. Regarding the force sensors, a similar observation, as when using two sensors combined, can be made. The force sensors $3, 5, 6$ and 7, which are located over the whole leg, are used. Here as well a tendency that some muscles play a more important role than others cannot be observed.

TABLE III: Classification accuracies using 2-4 sensors combined.

Amount Sensors	Force Sensor	IMU Sensor	Accuracy [∅%]
2	-	1,3	86.35
	7	3	85.34
	-	1,2	84.66
	5	3	84.60
	2	3	84.29
	3	3	83.86
	7	1	83.73
	5	1	83.42
	6	3	83.05
	1	3	83.03
3	5	1,3	89.53
	7	1,3	89.08
	6	1,3	89.04
	3	1,3	88.81
	2	1,3	88.42
	-	1,2,3	88.38
	1	1,3	88.21
	6	1,2	87.95
	5	1,2	87.90
	3	1,2	87.69
4	6,7	1,3	90.73
	3,6	1,3	90.69
	5,6	1,3	90.48
	3,5	1,3	90.44
	5,7	1,3	90.42
	1,5	1,3	90.33
	6	1,2,3	90.30
	5	1,2,3	90.26
	2,6	1,3	90.21
	2,5	1,3	90.17

For the other combination results when using 2-4 sensors combined, which are not presented here in detail, the IMU sensor 1 and 3 achieve better results than IMU sensor 2 or no IMU sensor. Force sensor 1 and 2 become more relevant. Using more sensors the differences between the worst and best results gets smaller. When IMU and force sensors are used in combination the difference in classification accuracy is at least for 2 used sensors $45\,\%$, for 3 used sensors $32\,\%$ and for 4 used sensors $23\,\%$. The difference gets smaller with increasing amount of sensors.

When using 5-10 sensors combined there is first a small increase in the classification accuracy as listed in Table IV. Here, the differences between the sensor combination results become very small. Therefore, only the best sensor combinations for each dimension are listed. Table IV reveals that using up to 6 sensors combined the classification accuracy still increases. From 7 to 10 it decreases and increases slightly. An important observation at this point is that we already achieve with 6 sensors ($92.20\,\%$) about the same classification performance as when using all 10 sensors ($92.40\,\%$).

TABLE IV: Classification accuracies using 5-10 sensors combined.

Amount Sensors	Force Sensor	IMU Sensor	Accuracy [∅%]
5	3,5	1,2,3	91.16
6	1,3,5	1,2,3	92.20
7	1,5,6,7	1,2,3	92.61
8	1,2,5,6,7	1,2,3	92.85
9	1,2,3,5,6,7	1,2,3	92.77
10	1,2,3,4,5,6,7	1,2,3	92.40

E. Systematic Exploration

The analysis was also conducted with the systematic exploration of the features space approach described in Mandery et al. [19]. We want to compare if the same results of the *brute-force* approach can be achieved compared to the approach of [19]. If that is the case, future analysis can be faster executed since not all sensor combinations have to be tested. As in the *brute-force* approach, we used all 10 subjects, all 14 motion classes and a window size of $300\,\mathrm{ms}$. In the approach of [19] a stratified 3-fold cross validation instead of a stratified 5-fold cross validation was applied which we also chose. This leads as well to a reduced training and testing time.

The results of both approaches differ slightly due to the different k-fold cross validation values and amount of training steps. But the overall tendency of common used sensors is the same. Also here, due to the similar classification results per dimension, can not be spoken of a best sensor combination for dimensions 2 or more. Since a combination of only force sensors performs worse than using a combinations of the sensor modalities or only IMU sensors, for dimension 3 and higher there are no combinations of just using force sensors further trained. When just using one or two sensors combined, IMU sensor 2 performs much worse than the other two IMU sensors. Therefore, for dimensions 3 and higher, using only IMU sensor 2 is often not trained and tested due to the worse results before. Since these ones are more outliers and do not conduct after the common tendency, it is not drastic that these are thrown out by the approach of [19]. With our data and feature sets it was possible to reduce the training and testing time by one third with the systematic exploration of the feature space approach compared to the *brute-force* approach.

V. Conclusion

In this paper, we investigated the quality of our motion classification approach using Hidden Markov Models for an unilateral, lower limb passive exoskeleton which is equipped with seven 3D force sensors and three Inertial Measurement Units (IMUs) based on our previous work [2]. Our data set contains 14 different daily activities of 10 subjects. In our first evaluation, we compared the classification accuracies when all subjects (AS) or single-subjects (SS) are considered in the training set. Also for small window sizes, the classification accuracy drastically increases when training on SS.

In the two following evaluations we analysed in detail how each single motion performed in the AS and SS evaluation. The motions *Turn Left (TL), Turn Right (TR), Going Upstairs (GU), Going Downstairs (GD)* and *Going Downstairs Backwards (GDB)* performed in the SS evaluation slightly worse than the other motions. In the AS evaluation the differences between the single motion types were higher. Furthermore, similar motions such as *GD* and *Walking Forward (WF)*, *GDB* and *Walking Backwards (WB)*, *TR/TL* and *WF* were often mixed up in the classification process due to their similar locomotion elements.

Our final analysis covered the search of a minimalistic sensor setup. Here, we compared the results of a *brute-force* approach with a systematic exploration of the feature space and have shown that the overall tendencies of common used sensors remain the same and no important combinations are discarded. Using a combination of just 6 sensors we can already achieve about the same classification accuracy as using 10 sensors. With our systematic exploration we can reduce the training time by one third.

Our approach and sensor setup were applied so far to one specific exoskeleton. The same sensor positions should have the same results on other exoskeletons. We want to verify this in future studies with our new adjustable lower limb exoskeleton. Future tests should further be conducted with active exoskeletons to verify if the motion patterns change significantly.

We will conduct deeper analysis based on the results of the systematic exploration of the feature space approach of Mandery et al. [19]. Here, we want to investigate if the best sensor combination changes with the window size. Furthermore, we want to analyse if sensor combinations are subject-specific when training on single subjects. Additionally, we will investigate if derived features of our data set achieve better results.

Acknowledgement

The authors would like to acknowledge the help and support of our student Tilman Daab of the H^2T lab who contributed to Section IV-E by adapting our previous work to systematically explore the feature space and helped to identify the differences between the results of both methods.

References

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

[2] 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 International Conference on Intelligent Robots and Systems (IROS)*, pp. 5431–5436, IEEE, 2018.

[3] 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.

[4] 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 *4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pp. 1794–1800, 2012.

[5] 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 *35th annual international conference of the IEEE Engineering in medicine and biology society (EMBC)*, pp. 5606–5609, 2013.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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 *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, IEEE, 2008.

[13] 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.

[14] 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.

[15] 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.

[16] J. Taborri, E. Scalona, S. Rossi, E. Palermo, F. Patanè, and P. Cappa, "Real-time gait detection based on hidden markov model: is it possible to avoid training procedure?," in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 141–145, 2015.

[17] M. Wang, X. Wu, D. Liu, C. Wang, T. Zhang, and P. Wang, "A human motion prediction algorithm for non-binding lower extremity exoskeleton," in *IEEE International Conference on Information and Automation*, pp. 369–374, 2015.

[18] C. Gong, D. Xu, Z. Zhou, N. Vitiello, and Q. Wang, "Real-time on-board recognition of locomotion modes for an active pelvis orthosis," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, pp. 346–350, IEEE, 2018.

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

[20] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[21] C. Freeman, D. Kulić, and O. Basir, "An evaluation of classifier-specific filter measure performance for feature selection," *Pattern Recognition*, vol. 48, no. 5, pp. 1812–1826, 2015.

[22] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.