From Activity Recognition to Motion Assessment: Delimitate against the Other Class within a WBAN

Martin Seiffert
Dep. of Mathematics and Computer Science
Freie Universität Berlin
Takustraße 9, 14195 Berlin, Germany
seiffert@inf.fu-berlin.de

Flavio Holstein
Dep. of Computer Science
Technical University of Berlin
Straße des 17. Juni 135, 10623 Berlin, Germany
flavio.holstein@campus.tu-berlin.de

Jochen Schiller
Dep. of Mathematics and Computer Science
Freie Universität Berlin
Takustraße 9, 14195 Berlin, Germany
schiller@inf.fu-berlin.de

ABSTRACT

Wearable sensor nodes use activity recognition systems to identify human activities. However, several applications such as physical rehabilitation and professional sports coaching require not only the identification of motion but also quality assessment to provide appropriate feedback to the user. In this work, we present AAC, a generalized trainable process chain for the online assessment of periodic human activity within a WBAN. AAC evaluates the execution of separate movements of a prior trained activity on a fine-grained quality scale. We connect qualitative assessment with human knowledge by projecting the AAC on the hierarchical decomposition of motion performed by the human body as well as establishing the assessment on a kinematic evaluation of biomechanically distinct motion fragments. We evaluate AAC in a real-world setting and show that AAC delineates movements of correctly performed activity to faulty ones and provides detailed reasons for the activity assessment. Both are crucial for an appropriate feedback to the user.

CCS Concepts

•**Computer systems organization** → *Sensor networks;*
•**Computing methodologies** → *Instance-based learning; Anomaly detection; Distributed artificial intelligence;*

Keywords

activity assessment chain; motion quality; wearable sensors;

1. INTRODUCTION

Driven by cheap to produce and easy to apply ubiquitous wireless sensor nodes (hereafter nodes) equipped with inertial sensors such as accelerometers and gyroscopes, activity recognition (AR) based on wearable nodes is a field broadly explored in recent research. Such AR systems typically utilize the data of multiple nodes attached to the limbs of the human body to recognize a predefined set of common activities. To achieve this goal, machine learning algorithms are used to train a model from prior labeled training data of the wireless body area network (WBAN). Subsequently the model is used to classify new data of the WBAN as one of the prior trained activities.

In [4] Bulling et al. give a tutorial to the typically used general-purpose framework called activity recognition chain (ARC). The ARC comprises raw data acquisition of multiple nodes, preprocessing of the data to remove noise and interfering artifacts, data segmentation, feature extraction and selection, and finally activity classification and decision fusion. Many of the current AR approaches are focused on improving the basic processes of the ARC. In [11] Khan et al. investigate optimal sampling rates for accelerometry based AR to reduce system requirements related to energy consumption or memory whilst retaining recognition accuracy. In [2] Banos et al. investigate signal segmentation and present a study that analyzes the effects of the windowing process on AR system performance. In [7] Ghasemzadeh et al. present a feature selection approach that considers classification accuracy as well as the energy consumption required for feature computation.

More complex approaches focusing on generic activity recognition use the concept of hierarchical decomposition of activities into time series sub patterns. These approaches in general utilize processes of the ARC to classify low level sub patterns of activities first, to identify higher level activities as sequential or concurrent combinations of sub patterns. These approaches contribute to real-world problems, for example: AR on streaming data to identify activities of interest among other activities [13, 15], recognition of concurrent and interleaved activities [17, 16], training of AR systems based on limited training data [19], or sharing AR Systems across platforms [20].
Beyond the classification of activities, a WBAN equipped with inertial sensors contains great potential for qualitative activity analysis [5]. Feedback relating to the quality of execution of an activity is a very interesting field in AR especially within the application domains of sports and healthcare. In sports training and physical rehabilitation, feedback regarding wrong body movements is crucial for learning and improvement of motor skills and physical fitness [21, 6].

Most approaches to quality evaluation of activities are straightforward and implement expert knowledge of a specialized application domain [18, 23, 14]. However, these

BODYNETS 2016, December 15-16, Turin, Italy
Copyright © 2017 EAI
DOI 10.4108/eai.15-12-2016.2267648

assessment approaches are not designed to generalize well beyond the corresponding application domain. There are a number of approaches that handle activity assessment as a classification problem [25, 9, 1]. This research successfully demonstrates that AR can be used to detect erroneous activity executions, but the inherent requirement to train defective activities restricts these approaches to applications were this is affordable.

Research on more generalized activity assessment is very scarce. In [12] Khan et al. address the challenging problem of generality and propose a framework for skill assessment in the context of AR that enables automatic quality analysis of complex human activities without domain-specific knowledge as a part of the system. The framework utilizes a rule induction tree built on the symbolic abstraction of the raw sensor data of an accomplished activity. The quality measure is computed from features concerning the complexity of the rule tree as well as metrics of the induced transitions. The final rating of new activity samples is based on the ratings of prior trained activity samples in different quality levels. In [22] we introduce a trainable system architecture for complex online motion assessment that requires the training of the correctly executed activity only. The assessment is based on the classification of motion fragments as atomic parts of an intuitive description of a body movement. The system identifies and assesses movements of an activity by reference to the amount of correctly classified motion fragments. The presented approach states very good identification accuracy of motion fragments but lacks in definite differentiation between correct and faulty movements.

The contribution of this paper is three-fold: i) First we classify methods for quality quantification in AR what was not done before. We give reason for the method we use and show how semantic property's can be added. ii) We extend our work in [22] by investigating the delimitation of a trained activity against untrained activity as well as the detailed assessment of motion fragments. As a result, we present a generalized activity assessment chain (AAC) for the detailed assessment of cyclic activity, which is designed to work in a fully distributed fashion within a WBAN. iii) We implement AAC and evaluate the applicability and the practical value in a real-world case-study of indoor rowing.

2. ASSESSMENT METHOD

	DM	EML	IML
configuration effort	+	-	(+)
semantic properties	+	+	-
generalization ability	-	(+)	+

+ favourable (+) constrained - adverse

Figure 1: Comparison of assessment methods

For the assessment of activities, a reference for quality quantization is needed. We differentiate between three methods used in the related work to model this reference: domain modeling (DM), explicit machine learning (EML), and implicit machine learning (IML). As described below and depicted in Fig. 1, these methods differ in the effort necessary for the inertial system configuration (Configuration Effort), in the meaning of their computed assessment values (Semantic Properties) as well as in their ability to be transferred across multiple application domains (Generalization Ability).

Domain Model. The DM method relies on manually modelling domain specific features based on expert knowledge [23, 18, 14]. In [14] for example, the quality of climbing moves is determined based on features implemented specifically for rock climbing. These approaches have the advantage that, based on the semantic expert knowledge included in the implemented domain specific assessment model, feedback concerning the activity assessment can be easily connected to the knowledge of the user. In addition, after implementation, these approaches do not necessarily need an inertial training, neither of correctly nor of incorrectly conducted samples of the activity. Hence, no inertial configuration effort is necessary. However, domain specific modelling do not generalize well, as the implemented domain specific logic cannot be easily applied to other application domains.

Explicit Machine Learning. For the EML approach, an explicit model for any quality class which should be distinguished is learned by a supervised training of a classifier on appropriately conducted activity samples. Activity assessment is thus translated into a classification task as it is done in [25, 9, 1]. Subsequent new samples of the activity are assigned to one of the previously trained quality classes. As a result of the supervised training, the domain specific meaning of the different quality classes is known. Thus, an appropriate semantic feedback to the activity assessment is possible. Systems using this approach can be transferred to other application domains where an appropriate training of correctly as well as incorrectly conducted activity samples is possible. Depending on the intended granularity of the activity assessment, the inertial training of the quality classes can be expensive. Nevertheless, the training of all intended quality classes is not always possible. Especially in motor learning or physiotherapy applications, the accomplishment of the purposefully incorrectly performed activity is not recommended. In this case the approach is restricted to applications where historical data on incorrectly performed samples of the activity is available. Another drawback of the EML approach is that the resolution of the assessment is limited to the amount of explicit trained quality classes. [12] addresses this by assigning relative quality labels through pairwise comparison of explicit trained activity samples by a domain expert and thus providing a quality scale. New activity samples are assessed by ranking them within the prior trained quality scale.

Implicit Machine Learning. By IML we summarise approaches where a model is generated from ground truth data representing activity samples at the best affordable quality only [22, 3, 24, 8]. Upcoming activity samples are assessed by delimitation of the previously modelled ideal quality class against motion data which do not correspond to the trained activity. Hence, this approach makes use of algorithms known from the *null class* problem [4], also known as the problem of the *other class* [13] respectively. This can be seen as a form of implicit learning, in other words, learning about faulty conduction of the activity without knowing the faults in particular. In [3], for example, selected feature values, which were extracted from the raw data of a movement, are checked for predetermined thresholds. If a feature value exceeds the corresponding threshold, the movement is reported as faulty. As for the second approach, a supervised training is needed for generalization across application domains. But the training is limited to the correct conduction of samples of the intended activity only, which results in a reduced inertial configuration effort. As the supervised training did not comprise any faulty activity samples, this method did not include semantic knowledge corresponding to faulty conductions by default.

In our work, we focus on proving a generalized process chain for activity assessment. Hence, the IML method can be generalized across multiple application domains by an affordable need of inertial training we make use of this method. As a drawback, this approach lacks in semantic knowledge, included by design (e.g. what is the assessment composed of and what can be done to improve the performance). We address this gap by adding biomechanically modelled aspects to the processes chain. As in [22] we establish the activity assessment on biomechanically distinct motion fragments and aggregate the motion assessments on different abstraction layers concerning the hierarchical decomposition of a motion performed by the human body. Thus we add temporal and spatial context to the assessment. In contrast to [22] we realize the fragment assessment by detailed evaluation of the feature values which characterize a motion fragment. Therefore we implement a descriptive set of kinematic features that give intuitive reason to the assessment of a fragment.

3. DESIGN CONCEPT: AAC

For the implementation of the IML method by consideration of the mentioned semantic demands we propose the concept of AAC for generalized activity assessment within body area networks as depicted in Fig. 2a. The AAC consists of a sequential flow of processes which are assigned to abstraction layers related to the hierarchical decomposition of motion performed by the human body, namely: *raw data layer*, *fragment layer*, *body part layer*, and *body layer*. The AAC contains processes known from AR which are marked with a grey background. We distribute the work load within the WBAN by processing the motion data of a certain body part concerning the first three layers directly on a node which is attached to the respective limb. On the last layer, the data computed on the nodes attached to multiple body parts are exchanged and fused to achieve the whole picture of the body movement.

Each abstraction layer contains processes of three data flows which accomplish different functions within the AAC: *recognition*, *reasoning*, and *assessment*.

Recognition. The recognition flow involves processes aiming for the basic identification of the trained activity and implements a typical ARC. On the raw data layer, the *segmentation* process recognizes biomechanically distinct motion segments within the processed raw data stream that contain information about a potential activity. Typically, before segmentation, the raw data passes a *filter* process to reduce noise and interfering information on the raw data signal. On the fragment layer, recognized segments are identified and labelled by the *fragment classification* process based on features extracted by a *feature extraction* process. On the body part movement layer in the resulting temporal sequence of classified motion fragments, a *sequential pattern mining* process identifies accomplished body part movements of the conducted activity. On the body layer, several concurrently identified body part movements are collected and combined by the *concurrent pattern mining* process to obtain the body movement in its entirety.

Reasoning. In a real-world setting not all parts of a continuous data stream are relevant for the trained activity, e.g. when disrupting the trained activity with other activities. To avoid diffusing assessments, the reasoning flow discards motion data which do not correspond to the trained activity and thus handles the problem of the other class on multiple layers. On each abstraction layer, a *delimitation* process verifies the decision of the preceding recognition process and thus determines which motion data will be handled by the subsequent *assessment* process and which motion data will be discarded. Depending on the algorithm utilized for recognition, the delimitation process can be included in the preceding recognition process.

Assessment. The assessment of biomechanically distinct motion fragments passing the delimitation process on the motion fragment layer ensures that the assessment can be intuitively connected to natural parts of movements of an activity (see fig. 2 b)). The fragment quality (QF) can be derived on the detailed evaluations (QX) of the features extracted by the feature extraction process. The quality of recognized body part movements (QP) as well as the body quality (QB) can be aggregated from the assessments of the respective previous abstraction layer. Assessments on the fragment layer contain temporal semantic as they are connected to successive motion fragments. Assessments on the body part layer contain spatial semantic by their assignment to body parts.

WBAN. The AAC is designed to run online and completely distributed within a WBAN. Hence we respect the limited resources of wearable wireless nodes: computational power, memory, and energy. With the abstraction of the raw data by feature extraction and numerical classification early at the fragment layer, memory needed for the expensive raw data buffering is bounded to the size of a single motion fragment. Additionally, the complexity of the data processing at higher abstraction layers is reduced to save computational power at subsequent processes of the AAC. Up to the body part layer, no data had to be exchanged between the nodes. For the body layer, only a few data characterizing body part movements are exchanged between nodes to support recognition, reasoning, and assessment (e.g. time stamp, duration, and quality). Thus we respect the energy consumption needed for costly wireless communication.

4. IMPLEMENTATION: AAC

As almost all movements of the human body are made possible by the joints and are mainly rotational, we focus on angular motion data gathered from a three-axis gyroscope. We apply a Butterworth low-pass filter of second order with a cutoff frequency of 3 Hz for the reduction of noise and interfering information within the angular motion signal. For segmentation, we stick with the biomechanic segmentation algorithm in [22] as previous results demonstrate that this approach reliably produces segments of angular raw data based on kinematics of human motion.

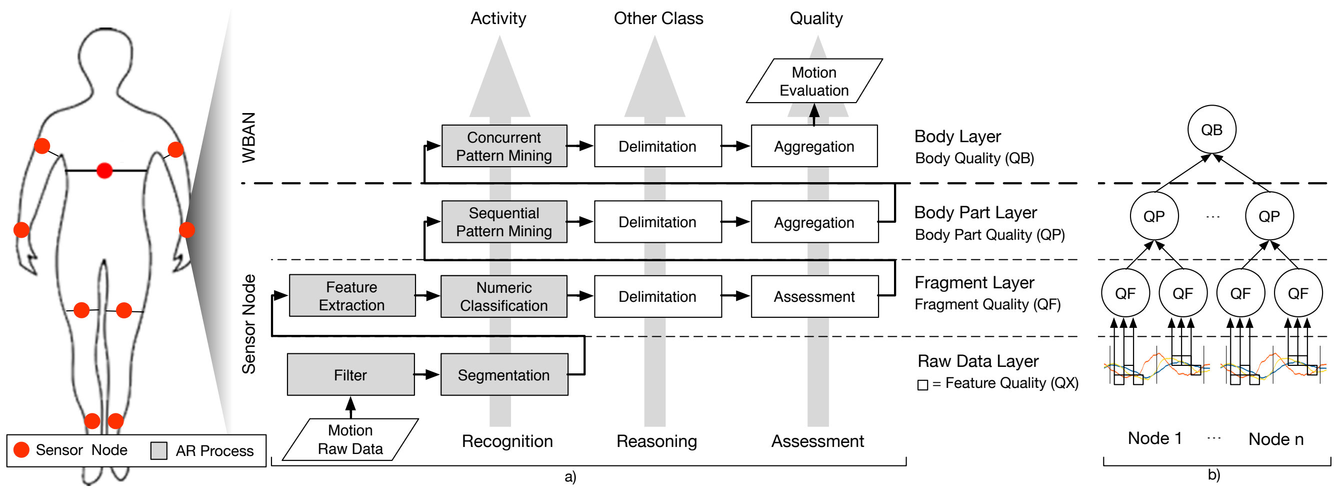

Figure 2: a) AAC b) Quality aggregation model

Fragment Layer. Within the feature extraction process on the fragment layer, we use four kinematic feature-types [21] applied separately to the three axis of the angular motion data obtained from the gyroscope. We measure *displacement* (DP), *maximum velocity* (MV), *average velocity* (AV), and *time-to-peak* (TTP). We compute DP straightforwardly by computing the integral from the angular velocity signal. It is worth mentioning that motion segments built by the utilized biomechanical segmentation algorithm in general comprise small data sequences of a few seconds. Thus, the error generated during integration is tolerable. By TTP we measure the time from the beginning of the motion fragment until the maximum of the velocity signal is reached and divide the value by the length of the motion fragment. In total, we compute a feature vector $x = (x_1, x_2, ..., x_n)$ composed of $n = 12$ feature values from the raw data of any motion segment.

For the classification of motion segments we implement a classical nearest prototype classifier [10] characterized by a set W of prototypes $w \in \mathbb{R}^n$, with w representing the feature vector and $c(w) \in \{1, ..., Z\}$ its assigned class label. The class of a feature vector x of a new segment is defined by Eq. 1.

$$c(x) := c(arg \min_{w \in W} \{d(w, x)\}) \tag{1}$$

We normalize x before classification and utilize the Euclidean distance measure as distance measure d. As in [22], we derive W by k-means cluster analysis of feature vectors extracted from motion fragments of movements that where performed while the training of the activity. The number of partitions k is provided by the number of motion fragments expected for a movement of the trained activity.

For delimitation we consider the minimum bounding hyper rectangle R_c^T covering the set V_c^T of feature vectors v, which represents all motion fragments observed during the training (T) for the classification $c(x)$, see Eq. 2.

$$R_c^T := \prod_{i=1}^{n} [v_{i\wedge}^{T,c}, v_{i\vee}^{T,c}] \; with \; \begin{cases} v_{i\wedge}^{T,c} := \min\limits_{v \in V_c^T} v_i \\ v_{i\vee}^{T,c} := \max\limits_{v \in V_c^T} v_i \end{cases} \tag{2}$$

The training contains only examples of the correctly performed activity. Hence we add a margin to the intervals of R_c^T so that motion fragments can pass the delimitation process even if they contain a limited error in their execution. This results in an extended (E) hyper rectangle R_c^E, see Eq. 3, $\delta > 0$.

$$R_c^E := \prod_{i=1}^{n} [v_{i\wedge}^{E,c}, v_{i\vee}^{E,c}] \; with \; \begin{cases} v_{i\wedge}^{E,c} := v_{i\wedge}^{T,c} - \delta * (\overline{v}_i^{T,c} - v_{i\wedge}^{T,c}) \\ v_{i\vee}^{E,c} := v_{i\vee}^{T,c} + \delta * (v_{i\vee}^{T,c} - \overline{v}_i^{T,c}) \\ \overline{v}_i^{T,c} = \operatorname{mean}\limits_{v \in V_c^T} v_i \end{cases} \tag{3}$$

We define that any feature vector x outside of R_c^E did not represent the classified motion fragment of the trained activity and belongs to the other class. Hence we relax the classification and reject x and the motion segment respectively. Any motion segment which passes the delimitation process is regarded as a fragment of a movement of the trained activity and assessed by focusing on the position of x in the decision space in detail. Depending on $c(x)$ we compute the quality q_c^X for each feature x_i by quantizing the error with a linear function crossing the value range between R_c^T and R_c^E, see Eq. 4, Fig. 3.

$$q_c^X(x_i) = \begin{cases} 1 - (v_{i\wedge}^{T,c} - x_i)/(v_{i\wedge}^{T,c} - v_{i\wedge}^{E,c}) & \text{if } x_i < v_{i\wedge}^{T,c} \\ 1 - (x_i - v_{i\vee}^{T,c})/(v_{i\vee}^{E,c} - v_{i\vee}^{T,c}) & \text{if } x_i > v_{i\vee}^{T,c} \\ 1 & \text{otherwise} \end{cases} \tag{4}$$

We determine the fragment quality q^F as the mean of the single feature qualities of x. The scaling parameter δ defines the tolerance of error accepted for the motion fragment layer as well as the value range for quality quantization.

Body Part Layer. On the body part layer in sequence recognized fragments are represented by $f = (c, t_s^F, t_e^F, q^F)$ where c represents the class label, t_s^F the start-time, t_e^F the end-time, and q^F the assigned fragment quality. The sequence of classifications is compared to a reference symbol sequence which represents the sequence of classifications determined for the trained activity at the dedicated body part

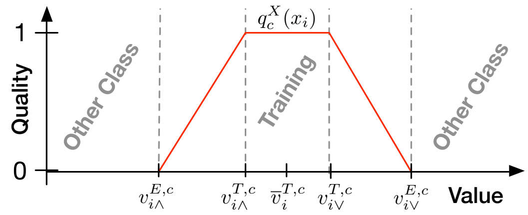

Figure 3: Quality Assessment Model

out of the training data, as in [22]. As distance measure we use the Levenshtein distance which is a well-known metric in AR [16, 17, 22]. If the measure is below a predefined threshold γ, a body part motion is recognized. A body part quality q^P is assigned by the mean of the qualities of the motion fragments which the body part motion consists of. Fragments which do not belong to any match are considered as motion of the other class and thus rejected.

Body Layer. On the body layer in parallel recognized body part movements of several nodes arrive, $m = (id, t_s^P, t_e^P, q^P)$, and are stored in a buffer M of size r, where r is determined by the number of nodes attached to the body, id is the node id, t_s^P the start time, and t_e^P the end time of the body part movement. For recognition of the overall body movement the relative temporal coverage $u(M)$ of buffered body part movements is determined by Eq. 5.

$$u(M) = (\min_{m \in M} t_e^P - \max_{m \in M} t_s^P)/(\max_{m \in M} t_e^P - \min_{m \in M} t_s^P) \qquad (5)$$

Given a minimum coverage threshold p^T, a body movement is recognized if $u(M) \geq p^T$. p^T is defined by the minimum coverage determined for a movement of the training data $u_{min}^T * \epsilon, 0 < \epsilon \leq 1$. Incoming body part movements supersede the body part movement $m \in M$ which ends first. Body part motions which are not part of a recognized body movement are rejected. The body quality q^B for a recognized body movement is computed by the mean of the qualities of the respective body part movements.

The triple δ, γ, and ϵ controls the sensitivity of the reasoning flow of the AAC. A small δ, small γ, and big ϵ induce a sensitive delimitation on the abstractions layers closely to the trained activity. A bigger δ, bigger γ, and smaller ϵ induce more tolerance for error within the evaluated activity.

5. CASE STUDY: INDOOR ROWING

For evaluation, we chose the assessment of rowing activity on an indoor rowing machine. Rowing is a continuous motion comprising a sequence of strokes. The rowing cycle at an indoor rowing machine comprises two major phases: *drive* and *recovery*. Regarding this, two positions are relevant: *catch* and *finish*. The drive phase starts with the catch of the handle and is initiated with a push from the legs. After the legs are almost extended, the rower first begins to lean back and then the arms begin to draw the handle toward the body until the stroke is finished. The recovery phase begins at the finish and is initiated with the arms straightening followed by the trunk raising. If the arms are almost straight and the trunk reaches the upright position, the legs begin to flex until the shins are vertical and the catch position is reached again.

5.1 Setup

Driven by the needs of the application domain for distributed activity assessment within WBANs, we make use of the sensor board F4VI2 [26]. For the experiment we use the data of the three-axis gyroscope of the integrated nine degrees of freedom motion tracking device MPU9150 at a sampling frequency of 100Hz. For characterization of the rowing activity we use three F4VI2 nodes attached to the human body (see Fig. 4). We add one node to the right wrist (RW), one node to the upper back (UB), and one node to the upper right leg (RL). The rowing activity was performed on a concept 2® indoor rowing machine of type D. The nodes are synchronized by broadcast at the start-up sequence.

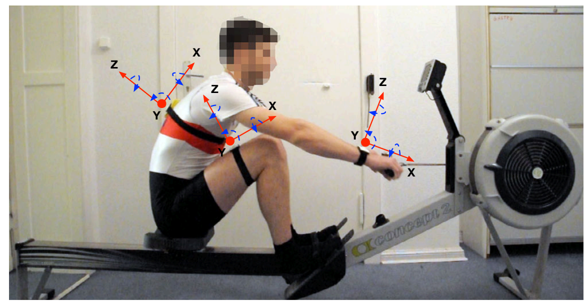

Figure 4: Setup and node placement

We record rowing activity of two rowing athletes $RA1$ and $RA2$ with several years of rowing experience. The rowing activity was supervised by an expert with more than ten years of rowing experience and documented by a video camera. All subjects are instructed by the expert to perform a continuous sequence of rowing activity in six different qualities C and $E1$-$E5$. C represents correctly performed rowing strokes while $E1$ to $E5$ include a certain technical error.

E1: At the finish the rower did not lean back far enough. The stroke did not reach the maximum power.
E2: The rower leaned back too far. The energy costs for leaning back too far are greater than the gain from rowing a longer stroke.
E3: The rower accelerated the handle too much at the finish. Instead of pulling the handle toward the body, he pulled himself forward toward the handle.
E4: The rower did not complete the pull of the handle toward the body and wasted a few centimeters of the stroke.
E5: The rower rushed down the slide during the recovery. He accelerated and wasted energy to stop the movement at the catch.

Every subject starts the rowing session with two sequences of 22 correctly performed strokes. Thereafter the subject performs five sequences of 22 technically incorrect strokes according to the quality classes $E1$-$E5$. In total, every subject performs 154 strokes. The recorded motion data contain periods of resting between rowing sequences as well as the period of attaching the nodes at the beginning of the session. In total, every rowing session resulted in a data set of about 17 minutes. We train the AAC for every subject separately on the first sequence of 22 correct performed strokes. We skip the first and the last stroke, as these are influenced by taking and putting back the handle respectively, and train the AAC on the remaining 20 strokes. Therefore we provide $k = 2$ equal to the number of phases of a rowing stroke: drive and recovery. For activity recognition and assessment, the rest of the session including 132 strokes is processed by the trained AAC. We provide $\delta = 50$, $\gamma = 0$, and $\epsilon = 0.6$, as this tolerance configuration recognizes correctly conducted as well as faulty strokes of the rowing activity while other motion is rejected very well.

5.2 Results

First we evaluate the recognition and delimitation flow of the AAC followed by focusing on the assessment of the rowing activity. In detail for RA1 and RA2 and all nodes together 2137 fragments are recognized on the fragment level. Through cooperation of the nodes, $31,5\%$ of these fragments are rejected on certain abstraction levels: $7,7\%$ on the fragment level, $13,1\%$ on body part level and $10,7\%$ on the body level. The AAC recognized $92,4\%$ of the 264 accomplished strokes on the body level. Every detected stroke belongs to a stroke sequence. The AAC missed 20 accomplished strokes as they differ too much from the trained strokes. Almost all belong to the start or the end of a stroke sequence. One stroke of RA2 within E5 could not be detected. The AAC is able to recognize correctly conducted strokes as well as faulty ones, while rejecting any motion belonging to the other class. These results constitute a very good recognition and delimitation accuracy and provide a strong basis for activity assessment.

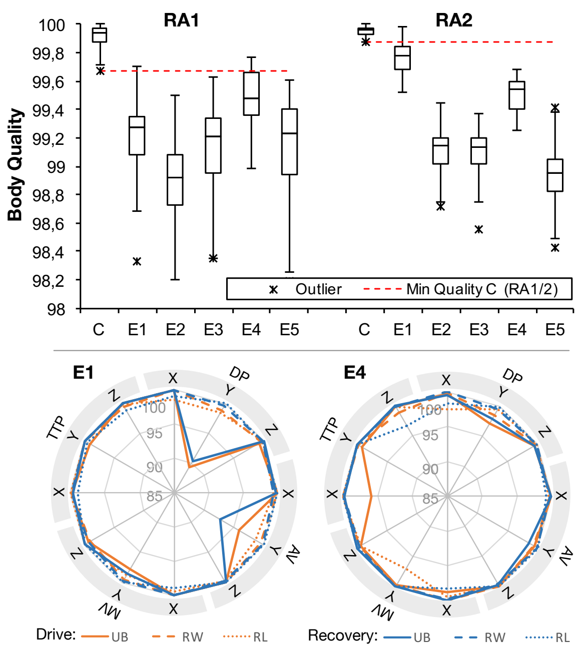

Figure 5: Body level assessment (top), detailed fragment level assessment (bottom)

Fig. 5 (top) depicts the final assessment of the 240 recognized rowing strokes on the body layer. The assessment shows a similar pattern of the ratings of the different quality classes of RA1 and RA2. As expected, the correctly performed strokes are rated with the highest quality and clearly stand out from the incorrectly performed strokes. $94,5\%$ of the incorrectly performed strokes are rated with a lower quality than the minimum quality reached by the respective correctly performed strokes. 6 strokes of E1 and 5 strokes of E4 are rated with a quality above the minimum quality of respective correctly performed strokes. The box plots for E1-E5 are comparatively tall, meaning a larger distribution of the assessments than for the correctly performed strokes.

The review of the video data, recorded while the rowing motion was carried out, states the distribution for E1-E5. Correct strokes are trained as usual for RA1 and RA2. Thus, rowing strokes at quality C at constant level is easier than the simulation of a certain technical error. In addition, the error in strokes of E1 and E4 includes only small deviations to strokes of respective quality C. The error of the back in E1 is estimated by the expert for RA2 in average to $6°$ and for RA1 in average to $14°$ compared to the back in the respective quality C. The error of the handle for RA1 in E4 amounts to $4cm$ in average to the respective quality C and thus causes only minimal deviations of the angular motion.

Furthermore we evaluate the temporal and spatial semantic in terms of which features of which motion fragment and which body part give reasons for the assessment. By reference to the aggregation model (Fig. 2b) we build 12-D spider plots for the visualization of the feature assessments. In Fig. 5 (bottom) the spider plots for E1 and E4 of RA1 are depicted. Therefore we compute the average of the feature assessments of E1 and E4 respectively. E1 is dominated by an error of DP-Y as well as AV-Y of the back within the drive and the recovery respectively. As E1 is defined by an angular error of the back at the finish and the rotation on the y-axis of UB correlates to the rotation of the back on the sagital plane of the human body, the assessment of E1 conducted by RA1 can be easily connected to the error description of E1. For E4, a correlation between the error description and the feature assessments is less evident. The worst rated features for E4 are TTP-X (95,9%) of UB followed by TTP-Z (96,8%) and MV-Y (97,0%) of RL. We expected to measure the worst error on RW because E4 is defined by an unfinished arm movement. The worst feature quality of RW is measured with DP-Y at the drive (98,6%).

The results of our case study demonstrate the practical value of the AAC as almost all strokes of the rowing activity are recognized, all motion of the other class is rejected and clear differentiation between rowing strokes in different qualities is possible. Furthermore, the AAC gives detailed insights concerning the reason of the activity assessment. For enough significant deviations of the conducted rowing activity from the trained activity, the AAC can provide a reasonable feedback which can be used to identify the cause of the error and improve the performance. The cause of error is not obvious for very small deviations.

6. CONCLUSION

In this paper we categorized the current solution space for quality quantification in AR. We found that most approaches are domain-specific and thus prevent re-use in different application scenarios. Hence we introduced and analyzed the concept of AAC, which is a generalized trainable activity assessment chain for distributed online evaluation of periodic activity within WBANs. Our activity assessment leverages the position of a characteristic set of kinematic features in the decision space to reject untrained motion and to achieve a fine-grained evaluation of distinct motion fragments. To include spatial and temporal semantics, AAC first decomposes the human motion and then aggregates multiple levels of the resulting hierarchical structure. Our case study shows that AAC can be applied with minor configuration effort to indoor rowing activity and is not only able to clearly distinguish rowing activity from other motion but also to provide a detailed reasonable assessment of rowing strokes in different qualities. Thus, an informational feedback to the user concerning the cause of the error becomes possible. In future work, we will compare AAC to additional scenarios. Furthermore, we will extend AAC by weighting feature assessments depending on the intensity of motion within the WBAN to work out the most relevant assessments.

7. REFERENCES

[1] R. Adelsberger and G. Troster. Experts lift differently: Classification of weight-lifting athletes. In *2013 IEEE International Conference on Body Sensor Networks (BSN)*, Cambridge, MA, USA, 2013.

[2] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas. Window Size Impact in Human Activity Recognition. *Sensors*, 14(4):6474–6499, 2014.

[3] G. Bleser, D. Steffen, M. Weber, G. Hendeby, D. Stricker, L. Fradet, F. Marin, N. Ville, and F. Carré. A personalized exercise trainer for the elderly. *Journal of ambient intelligence and smart environments*, 5(6):547–562, 2013.

[4] A. Bulling, U. Blanke, and B. Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput. Surv. (CSUR)*, 46(3):33, 2014.

[5] A. I. Cuesta-Vargas, A. Galán-Mercant, and J. M. Williams. The use of inertial sensors system for human motion analysis. *Physical Therapy Reviews*, 15(6):462–473, 2010.

[6] A. I. Cuesta-Vargas and J. Williams. Inertial sensor real-time feedback enhances the learning of cervical spine manipulation: a prospective study. *BMC Medical Education*, 14:120, 2014.

[7] H. Ghasemzadeh, N. Amini, R. Saeedi, and M. Sarrafzadeh. Power-Aware Computing in Wearable Sensor Networks: An Optimal Feature Selection. *IEEE Transactions on Mobile Computing*, 14(4):800–812, 2015.

[8] H. Ghasemzadeh and R. Jafari. Coordination analysis of human movements with body sensor networks: A signal processing model to evaluate baseball swings. *IEEE Sensors Journal*, 11:603–610, 2011.

[9] H. Ghasemzadeh, V. Loseu, E. Guenterberg, and R. Jafari. Sport training using body sensor networks: A statistical approach to measure wrist rotation for golf swing. In *Fourth International Conference on Body Area Networks (BodyNets 09)*, 2009.

[10] A. B. Graf, O. Bousquet, G. Rätsch, and B. Schölkopf. Prototype classification: Insights from machine learning. *Neural computation*, 21(1):272–300, 2009.

[11] A. Khan, N. Hammerla, S. Mellor, and T. Plötz. Optimising sampling rates for accelerometer-based human activity recognition. *Pattern Recognition Letters*, 73:33–40, 2016.

[12] A. Khan, S. Mellor, E. Berlin, R. Thompson, R. McNaney, P. Olivier, and T. Plötz. Beyond activity recognition: skill assessment from accelerometer data. In *2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2015.

[13] N. C. Krishnan and D. J. Cook. Activity recognition on streaming sensor data. *Pervasive and Mobile Computing*, 10:138–154, 2014.

[14] C. Ladha, N. Y. Hammerla, P. Olivier, and T. Plötz. ClimbAX: skill assessment for climbing enthusiasts. In *2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2013.

[15] R. J. M. Lemmens, Y. J. M. Janssen-Potten, A. A. A. Timmermans, R. J. E. M. Smeets, and H. A. M. Seelen. Recognizing Complex Upper Extremity Activities Using Body Worn Sensors. *PLoS ONE*, 10(3):e0118642, 2015.

[16] L. Liu, Y. Peng, S. Wang, M. Liu, and Z. Huang. Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors. *Information Sciences*, 340-341:41–57, 2016.

[17] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum. From action to activity: Sensor-based activity recognition. *Neurocomputing*, 181:108–115, 2016.

[18] R. Mooney, G. Corley, A. Godfrey, L. Quinlan, and G. ÓLaighin. Inertial sensor technology for elite swimming performance analysis: A aystematic review. *Sensors*, 16(1):18, 2015.

[19] L. T. Nguyen, M. Zeng, P. Tague, and J. Zhang. Recognizing new activities with limited training data. In *2015 ACM International Symposium on Wearable Computers*, 2015.

[20] D. Roggen, S. Magnenat, M. Waibel, and G. Tröster. Wearable Computing: Designing and Sharing Activity-Recognition Systems Across Platforms. *IEEE Robotics & Automation Magazine*, 18(2):83–95, 2011.

[21] D. J. Rose and R. W. Christina. *A multilevel approach to the study of motor control and learning*. Allyn and Bacon, Boston, 1997.

[22] M. Seiffert, N. Dziengel, M. Ziegert, R. Kerz, and J. Schiller. Towards motion characterization and assessment within a wireless body Area Network. In *8th International Conference on Internet and Distributed Computing Systems*, 2015.

[23] R. Thompson, I. Kyriazakis, A. Holden, P. Olivier, and T. Plötz. Dancing with horses: automated quality feedback for dressage riders. In *2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2015.

[24] A. A. A. Timmermans, H. A. M. Seelen, R. P. J. Geers, P. K. Saini, S. Winter, J. te Vrugt, and H. Kingma. Sensor-Based Arm Skill Training in Chronic Stroke Patients: Results on Treatment Outcome, Patient Motivation, and System Usability. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):284–292, 2010.

[25] A. Yurtman and B. Barshan. Automated evaluation of physical therapy exercises using multi-template dynamic time warping on wearable sensor signals. *Computer Methods and Programs in Biomedicine*, 117(2):189–207, 2014.

[26] M. Ziegert, N. Dziengel, M. Seiffert, S. Pfeiffer, and J. Schiller. A developer and a reference board for distributed motion evaluation in Wireless Sensor Networks. In *IEEE International Conference on Industrial Technology (ICIT '15)*, 2015.