An Open, Labeled Dataset for Analysis and Assessment of Human Motion

Andre Ebert, Chadly Marouane, Christian Ungnadner, and Adrian Klein

Ludwig-Maximilians-University
Institute for Computer Science
Mobile and Distributed Systems Group
Oettingenstrasse 67, 80538 Munich, Germany
andre.ebert@ifi.lmu.de

Abstract. Analysis of human activity, e.g., by tracking and analyzing motion information or vital signs became lots of attention in medical as well as athletic appliances during the last years. Nonetheless, comprehensive and labeled datasets containing human motion information are only sparsely accessible to the public. Especially qualitatively labeled datasets are rare, although they are of great value for the development of concepts concerning qualitative motion assessment, e.g., to avoid injuries during athletic workouts or to optimize a training's success.
Therefore, we provide an open and qualitative as well as quantitative labeled dataset containing acceleration and rotation data of 8 different body weight exercises, conducted by 26 study participants. It encompasses more than 11,000 exercise repetitions of which we extracted 7,352 into individual segments. We believe, that due to its structure and labeling our work is suitable to serve for development, benchmarking, and validation of new concepts for human activity recognition and qualitative motion assessment[1].

Key words: machine learning, activity recognition, motion assessment

1 Introduction

Automated monitoring and analysis of human motion by using motion sensors gained great attention during the last years. Reasons for that are the ubiquitous availability of powerful, portable, and small computing devices, e.g., smartphones, as well as the distribution of robust and cheap sensor platforms capable of motion tracking. This enables the analysis of complex and also spatially distributed human motion without being bound to a fixed area or tracking system, e.g., in contrast to the usage of cameras for visual analysis approaches. An important field for automated monitoring and motion assessment are medical appliances as well as the athletic context and physical exercises. Apps like

[1] Publication notes: The dataset will be published at `github.com/andrebert/body-weight-exercises` together with this paper's presentation on the MobiHealth conference 2017, taking place in Vienna, 14-16 November.

Freeletics[2] provide detailed instructions for challenging workouts and exercises as well as features for exercise tracking, such as counting of repetitions or distances. But these challenging exercises are often conducted by amateur athletes without being monitored or advised by professional instructors. Unfortunatly, incorrect execution of physical exercises may not only lead to less successful training results, it can also lead to serious harm and injuries [4, 1, 6]. We believe, that automated monitoring, analysis and assessment of human motion will enable the development of proactive feedback systems, which are capable of reducing such injuries and optimize training results. Moreover, the assessment of human motion is also of use within other areas, e.g., in a medical context to monitor ambulant patients, for gait analysis, for workflow optimization, and others.

To address this challenge, we developed a distributed sensor system called SensX [3]. Subsequently, we designed a study for our concept's evaluation and recorded more than 11,000 individual repetitions of 8 different body weight exercises. In our context, body weight exercises are defined as physical exercises which are conducted only with an athlete's own body weight and without the use of artificial training equipment, e.g., Sit-ups or Squats [5]. Moreover, we developed different approaches for activity recognition and generic qualitative assessment of human motion in [2].

Besides our own concepts concerning human motion analysis, there may be lots of new approaches and ideas of other researchers which can be tested and evaluated by using our dataset as a basis. Therefore, we publish all data recorded during our study, which is publicly available at GitHub[3]. Besides the raw data of all sensors, we also extracted single exercises into individual segments of adaptive length and labeled them regarding their class of quality as well as regarding their type of exercise. This paper provides an overview concerning the used sensor system, the data structure, and the describing meta data and is meant to function as a manual for the provided dataset.

2 Dataset

All in all, we recorded 11,087 individual repetitions of 8 different body weight exercises, of which we were able to extract 7,352 repetitions into individual segments of adaptive length. Subsequently, all details concerning the study design, the sensor system, the recording process itself, and the data structure are presented.

2.1 Study design

In context of our study we recorded data of 26 athletes by using the SensX sensor system, which is presented in [3]. In prior to the collection of data we designed a workout plan encompassing 8 body weight exercises, scheduled to stress an

[2] http://www.freeletics.com
[3] https://github.com/andrebert/body-weight-exercises

athlete's body consistently. An overview across these exercises is depicted in Figure 1. Namely these were (1) Crunches (cr), (2) Lunges (lu), (3) Jumping Jack (ha), (4) Bicycle Crunch (bi), (5) Squat (kn), (6) Mountain Climber (mo), (7) Russian Twist (ru), and (8) Push-ups (li). The exercises are numbered in order of their execution during the study – their abbreviations are used for storing them in our file system and are originating from the German exercise terms. All athletes had to complete 3 sets containing 20 repetitions for each exercise. In between all sets, a mandatory break of 30 seconds was scheduled. All exercises were introduced in prior to their first execution with a professional instruction video. All exercises were taped on video during their execution. Additionally, the participants were urged to fill out a questionnaire containing several questions concerning their age, their profession and their habits regarding the conduction of physical exercises. Moreover, we captured the amount of repetitions which each individual athlete was able to execute and encouraged all participants to make a subjective rating concerning the overall quality of their exercise sets within a range of 1 to 5. Thereby, 1 symbolized *very good*, 2 *good*, 3 *medium*, 4 *bad*, and 5 *very bad* in terms of quality.

2.2 Study participants

Within the scope of this study we recorded the acceleration and rotation data of 6 female and 18 male participants. Their average age was 27.15 years while the average weight of all participants was $67.04kg$. The oldest participant was 53 years old, while the youngest was 20. All participants were coerced to provide information concerning the individual frequency with which they are doing sports as well as a self-assessment regarding their level of fitness. The rating scale for both, frequency and level of fitness, ranged from 1 to 5. Corresponding to the rating's range in Section 2.1, 1 symbolizes *never* or *very low* while 5 symbolizes *every day* or *very high*.

2.3 Sensor system

The sensor system used for tracking the athletes consists of 4 external sensor platforms which are attached to the athlete's wrists and ankles and an additional central computation unit (CCU), e.g., a smartphone, which is attached to the participant's chest. For fastening the CCU, a common GoPro harness (see Figure 1) was used. All devices are tracking acceleration as well as rotation information in X-, Y-, and Z-dimension, respectively. This makes an overall count of 30 individual signals which are available to describe an athlete's movements. The CCU is used to 1) gather the signals of all connected external sensor units and 2) to store them with a synchronized timestamp on its internal storage together with its own sensor data. Within the studies conduction, we used two different CCU setups: the first two athletes were tracked within a preliminary test cycle by using a LG Nexus 5x running Android 5.0 Lollipop, while the main study was conducted with a HTC One (M7) running Android 6.0.1 Marshmallow. Reason

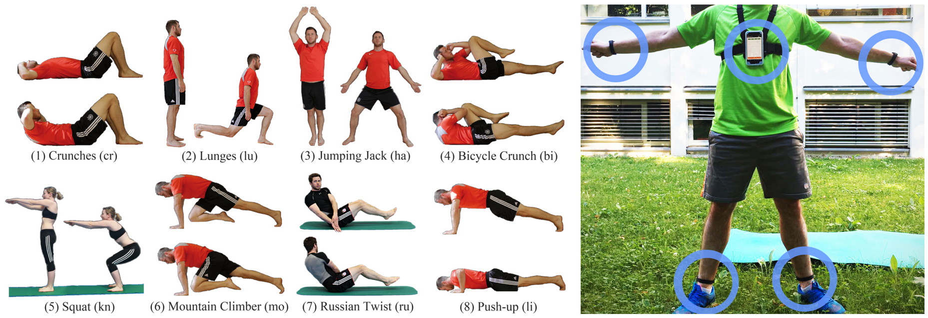

Fig. 1: All exercises which were included in our study plus their abbreviations (left) and the SensX sensor system carried by a study participant (right).

for that was, that the second setup was capable to record significantly more balanced sampling rates for the external sensors. Thereby, the LG setup operated with an average sampling rate of $200Hz$ for the CCU and an inconsistent rate of $20Hz$ - $40Hz$ for external devices, while the HTC setup achieved $100Hz$ for the CCU and a relatively consistent rate of $40Hz$ for external sensors. The main reason for that may be the manufacturer dependent implementation of Bluetooth Low Energy (BLE), which is used to connect the external sensors with the CCU. Our experiments showed, that different implementations offer different sampling rates when connecting multiple BLE devices at the same time.

2.4 Labeling

In general, the data set is labeled for two different purposes: 1) activity recognition and 2) qualitative motion assessment. All segmented repetitions are labeled implicitely by the folder structure used for storing their data files (see Section 2.6), which indicates the underlying type of activity (e.g., Sit-up, Push-up, Squat, etc.). A second, active labeling provides information concerning the actual quality of a conducted exercise and can be found within the `INFO` file, which is placed in the root folder of each participant's exercise sets. The qualitative labeling was implemented for the exercises 1-6 (see Figure 1) and in the frame of the qualitative assessment described in [2], while the exercises 7 and 8 are not labeled qualitatively. The second last row tagged with the `Index` key maps the quality ratings which are stored within the last row to specific exercise repetitions; the key used to identify the last row containing the ratings is named `Rating`. E.g., the rating related to the first Push-up within a set of Push-ups is the first value of the last row, while the rating of the second Push-up corresponds to the last row's second value (see also Section 2.6).

The qualitative labeling was undertaken by analyzing each single exercise repetition on basis of the recorded video material. To maximize the qualitative rating's accuracy, this step was done together with sports professionals especially skilled for body weight workouts and endurance training. At the beginning of the rating process, each individual exercise repetition is labeled with an initial start value of $p_s = 1$, which represents the highest quality (*very good*). Subsequently, a penalty value p_a is added to p_s for each anomaly a (mistakes, malpositions, etc.) made by the participant during an exercise's execution. If the characteristics of a specific anomaly are significant, p_a has a value of 1, for an anomaly of medium significance we add 0.5, and in case of a minor mistake it is 0.25. The amount of detected p_a was within a range of 7-13 incidents i, depending on the specific exercise. As soon as the score is reaching a value of 5, we stop adding any more penalty points, which means the rating scale's spectrum ranges from 1 till 5 (very good, good, medium, bad, very bad). In case that the final score is not an integer value, it became rounded commercially. Thus, the process of creating a qualitative label L for an individual exercise repetition r without rounding it may be defined as

$$L_r = p_s + \sum_{n=0}^{i} p_a, \text{ if } L_r > 5 : L_r = 5$$

Due to the sheer addition of penalty points without respecting specific causes we achieve a generic quality notation: L_r tells us about the quality of a specific exercise repetition r in terms of good or bad, though it does not contain any information concerning the cause for a specific rating.

2.5 Descriptive statistics

All in all we provide 7,352 adaptively segmented and qualitatively labeled exercise repetitions. Table 1 gives a descriptive overview across these quantities. Additionally, all raw data without qualitative labeling and segmentation can be found within each exercise set's root folder. It is significant, that a great number of the individually segmented repetitions were rated at least medium or even better. Reasons for that may be the relatively low average age of our participants as well as the fact that most of them do sports regularly and thus were not completely unskilled. Concerning Bicycle Crunches and Russian Twists, the dataset contains roughly half of the amount of repetitions (in contrast to all other exercises) – the reason for that is, that we decided to count a left-sided and the following right-sided execution together as one instance.

	cr	**lu**	**ha**	**bi**	**kn**	**mo**	**ru**	**li**
Overall quantity	1,315	1,345	1,394	692	1,379	1,227	-	-
Label *very good* (1)	554	592	235	150	458	439	-	-
Label *good* (2)	367	421	749	204	660	409	-	-
Label *medium* (3)	293	180	347	257	137	244	-	-
Label *bad* (4)	101	131	58	81	112	123	-	-
Label *very bad* (5)	-	21	5	-	12	12	-	-

Table 1: Qualitative rating and quantitative amounts of all segmented exercise repetitions.

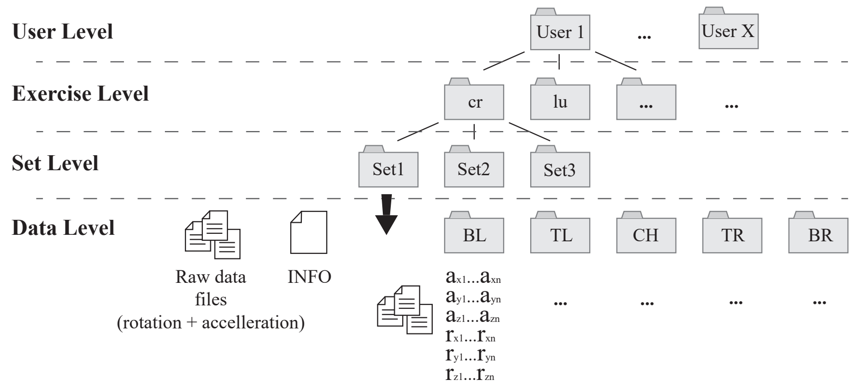

Fig. 2: The dataset's folder structure and the 4 different organization levels.

2.6 Folder and data structure

In the following we provide insights into the file and data storage system. The sensor data itself was stored by using the metrics which were returned by the specific sensor platforms. This means that data is stored in m/s^2 (minus earth gravitation) and rad/s for the CCU (Android API standard), while the external Metawear sensors return their measurements in gravitational force (g) and deg/s (for acceleration and rotation). All in all, our dataset is organized in four levels namely the (1) user level, (2) the exercise level, (3) the set level and (4) the data level (see Figure 2). The first level encompasses the folders of all 26 athletes which participated on our study. These are named by an individual code consisting of two abbreviating characters. Within the second level the exercise folders named by their German abbreviation can be found: 1) Crunches (cr), 2) Lunges (lu), 3) Jumping Jack (ha), 4) Bicycle Crunch (bi), 5) Squat (kn), 6) Mountain Climber (mo), 7) Russian Twist (ru), and 8) Push-up (li). Within the following third level the data is structured by the three individual workout sets, which each study participant had to conduct for each exercise (e.g., *Set1, Set2*, etc.). Each of these set folders encompass an instance of the fourth level, namely the raw data files of each exercise set, the segmented repetitions, and an `INFO` file with some meta information. The raw data files are named by the first two chars of the recording sensor platform's MAC address plus an indicator for acceleration (*acc*) or rotation (*rot*). The raw data files of the CCU are labeled with *CH* (chest) plus the rotation or acceleration tag. Furthermore, each raw data file is mapped to a specific body extremity:

- `C6_acc, C6_rot`: raw motion data of the left arm (*top-left*, abbr. TL)
- `EE_acc, EE_rot`: raw motion data of the left leg (*bottom-left*, abbr. BL)
- `D1_acc, D1_rot`: raw motion data of the right arm (*top-right*, abbr. TR)
- `CF_acc, CF_rot`: raw motion data of the right leg (*top-right*, abbr. BR)
- `CH_acc, CH_rot`: raw motion data of the chest (abbr. CH)

Next to the `INFO` file and the raw data files the data level contains 5 containers which are inhabiting the segmented repetitions of each set. They are named by the abbreviations introduced above and indicate from which sensor position the encompassed data originates (TL, BL, etc.). The files inside a specific sensor folder are named by the *acc*-tag or *rot*-tag plus the dimension of its content (x, y, or z) and an index number of the specific repetition, e.g., `acc-x-1`. This means that one single repetition of an exercise consists of 30 different data sources, e.g., the 10th repetition of a set is described by the following files, mapped to the five individually tracked parts of a participants bodies:

- right arm: TR/`acc-x-10`, TR/`acc-y-10`, TR/`acc-z-10`, TR/`rot-x-10`, TR/`rot-y-10`, TR/`rot-z-10`
- right leg: BR/`acc-x-10`, BR/`acc-y-10`, BR/`acc-z-10`, BR/`rot-x-10`, BR/`rot-y-10`, BR/`rot-z-10`
- chest: CH/`acc-x-10`, CH/`acc-y-10`, CH/`acc-z-10`, CH/`rot-x-10`, CH/`rot-y-10`, CH/`rot-z-10`
- left arm: TL/`acc-x-10`, TL/`acc-y-10`, TL/`acc-z-10`, TL/`rot-x-10`, TL/`rot-y-10`, TL/`rot-z-10`
- left leg: BL/`acc-x-10`, BL/`acc-y-10`, BL/`acc-z-10`, BL/`rot-x-10`, BL/`rot-y-10`, BL/`rot-z-10`

Data structure of segmented exercise repetitions The data within the files of segmented repetitions, e.g., `acc-x-10`, is organized as follows: the first row contains a vector of acceleration or rotation information; the second row contains a corresponding timestamp for each value counting the milliseconds of a repetition from start to end (for metrics see Section 2.6).

Raw data structure As stated in Section 2.6, each exercise set is described by ten files of raw data. Within these files an individual number of data rows with 4 columns can be found. The first column contains its timestamp in UNIX epoch milliseconds, the following three columns contain acceleration or rotation information in X-, Y-, and Z-direction.

Data format within the INFO file The `INFO` file contains additional meta information concerning individual exercise sets. Besides the average sampling rate of each single sensor platform it contains the following information:

- the amount of *Raw* data repetitions
- the number of segmented *Repetitions*
- the *Length* of a set in seconds
- the *Index* for the *Rating* for each single segmented exercise.

All keys are separated with an ":" from their content. Ratings and their indices are separated with an "-" in between the individual values.

2.7 Metadata

In order to enable an easy distribution and findability of our dataset, we defined meta information corresponding to a recommendation for dataset markups published by Google[4].

2.8 Adaptive Segmentation

All adaptively segmented exercise repetitions were created by using the segmentation approach described in [2].

3 Conclusion

Within this paper, we present an open dataset encompassing 11,087 repetitions of 8 different body weight exercises in raw data format. Moreover, 7,352 of these exercises were extracted into segments of adaptive length and labeled qualitatively as well as quantitatively. Subsequently they were stored in a suitable data structure for further analysis. Some promising results concerning human activity recognition and qualitative assessment of human motion on basis of this dataset were published in [2] and encourage further analysis. Multiple new analysis methods, e.g., neural networks, were not used for dataset examination, yet. Moreover, a detailed anomaly detection and description of specific error classes as well as advances for decision certainty concerning activity recognition and qualitative assessment are still missing. By publishing this dataset, we hope to promote further advances within human motion analysis, e.g., to optimize motion sequences and to detect malpositions or injuries within medical as well as athletic appliances.

References

1. Brady, T.A., Cahill, B.R., Bodnar, L.M.: Weight training-related injuries in the high school athlete. The American journal of sports medicine 10(1), 1–5 (1982)
2. Ebert, A., Beck, M.T., Mattausch, A., Belzner, L., Popien, C.L.: Qualitative assessment of recurrent human motion. In: European Signal Processing Conference (EUSIPCO), 2017 25th. IEEE (2017)
3. Ebert, A., Kiermeier, M., Marouane, C., Linnhoff-Popien, C.: Sensx: About sensing and assessment of complex human motion. In: Networking, Sensing and Control (ICNSC), 2017 14th IEEE International Conference on. IEEE (2017)
4. Jones, B.H., Bovee, M.W., Harris, J.M., Cowan, D.N., et al.: Intrinsic risk factors for exercise-related injuries among male and female army trainees. The American Journal of Sports Medicine 21(5), 705–710 (1993)
5. Lauren, M., Joshua, C.: Fit ohne Geräte: Trainieren mit dem eigenen Körpergewicht. Riva (2011)
6. Leaf, J.R., Keating, J.L., Kolt, G.S.: Injury in the australian sport of calisthenics: a prospective study. Australian Journal of Physiotherapy 49(2), 123–130 (2003)

[4] https://developers.google.com/search/docs/data-types/datasets