Determining the Single Best Axis for Exercise Repetition Recognition and Counting with SmartWatches

Bobak Mortazavi, Mohammad Pourhomayoun, Gabriel Alsheikh,
Nabil Alshurafa, Sunghoon Ivan Lee, Majid Sarrafzadeh
Wireless Health Institute
Computer Science Department, UCLA
Los Angeles, CA 90095
Email:{bobakm, mpourhoma, gabe.alsheikh, nabil, silee, majid}@cs.ucla.edu

***Abstract*—Due to the exploding costs of chronic diseases stemming from physical inactivity, wearable sensor systems to enable remote, continuous monitoring of individuals has increased in popularity. Many research and commercial systems exist in order to track the activity levels of users from general daily motion to detailed movements. This work examines this problem from the space of smartwatches, using the Samsung Galaxy Gear, a commercial device containing an accelerometer and a gyroscope, to be used in recognizing physical activity. This work also shows the sensors and features necessary to enable such smartwatches to accurately count, in real-time, the repetitions of free-weight and body-weight exercises. The goal of this work is to try and select only the best single axis for each activity by extracting only the most informative activity-specific features, in order to minimize computational load and power consumption in repetition counting. The five activities are incorporated in a workout routine, and knowing this information, a random forest classifier is built with average area under the curve (AUC) of $.974$, with average accuracy of 93%, in cross validation to identify each repetition of a given exercise using all available sensors and AUC of $.950$ with accuracy of 89.9% using the single best axis for each activity alone. Adding a gyroscope with the accelerometer increased the average AUC from $.968$ to $.974$, increasing the accuracy of specific movements as much as 2%. Results show that, while a combination of accelerometer and gyroscope provide the strongest classification results, often times features extracted from a single, best axis are enough to accurately identify movements for a personal training routine, where that axis is often, but not always, an accelerometer axis.**

I. Introduction

Body-wearable sensors for personal health monitoring have become an important tool in solving health problems. Chronic illness, which affects 133 million Americans [1] resulting in a majority of the health care costs [2], often stems from physical inactivity [3]. Indeed, cardiovascular disease and diabetes, the two most common chronic diseases resulting from inactivity are the source of an increasing economic burden on the United States. For instance, cardiovascular disease is projected to cost up to $818 billion in direct medical costs and $276 billion in indirect (loss of productivity) costs by 2030 [4], while diabetes, estimated at $245 billion in 2012, will increase due to the estimated prevalence doubling by 2050 [5]. As a result, there is a dire need of a solution that addresses this problem, and wireless health systems, used to monitor the physical activity of wearers [6], [7], are now looking to wrist-worn platforms for tracking various forms of exercises. Aerobic and resistance training exercises have both been shown to help prevent and address these diseases [8]. In some cases, resistance training alone can be effective [9].

Wearable sensor systems provide a large number of monitoring applications for energy expenditure. They can range from using only a single sensor [10] for general energy to networks of sensors for detailed applications, like swimming actions [11]. Monitoring energy expenditure while performing general daily activity [12], [13] is the most common such application. Calculating activity intensity, and approximation to metabolic equivalent of tasks allows for estimation of actual energy expenditure [14], [15]. A review of such wearable accelerometers and activity energy expenditures shows popularity of such systems in promoting physical activity [16]–[18]. For each existing platform, many application specific solutions exist as well, from design of gameplay [19], [20] to light-to-moderate physical activity monitoring [21].

With the popularity of Fitbit, and the emergence of the smartwatch platform from companies like Sony, LG, and Samsung, wrist-worn sensor devices for exercise monitoring will become an increasingly important tool in personal health monitoring. In particular, exercise routines and repetitions can be counted in order to track a workout routine as well as determine the energy expenditure of individual movements. Indeed, mobile fitness coaches [22] for counting repetitions of exercises [23], [24] have become a growing topic of interest, including the selection of sensors and locations for tracking activities [25]. Mobile fitness coaching has covered the range of topics from quality of performing such sports actions [26] [27] to detection of the specific sports activity [28].

This paper introduces a framework for platform creation (e.g., accelerometer only system vs. accelerometer and gyroscope) and machine learning of some activities, which can be especially useful in the emerging market of smartwatches. By identifying the most informative activity-specific features, a system can be optimized to reduce the computational load as well as the power consumption through appropriate sensor selection. In particular, the decision to add an accelerometer and a gyroscope, such as the Samsung Galaxy Gear [29], or accelerometer only in Fitbit and Sony platforms [30] gives rise to the questions of what sensory devices are necessary in

978-1-4799-4959-5/14 $31.00 © 2014 IEEE
DOI 10.1109/BSN.2014.21

IEEE computer society

a single smartwatch platform to make it a successful exercise tracking device for software platforms, such as those available on the market by companies such as Focus [31] (who provided the desired workout routine for this work) as well as similar platforms in research fields [32]. This work will attempt to analyze the performance difference for five workout exercises using a Samsung Galaxy Gear smartwatch platform in order to count repetitions for use in a personal gym coach application, by leveraging the contextual information of a given workout routine to classify movements and use said classification as a repetition counter. It will analyze the difference between the accelerometer, the gyroscope, and the combination there of, as well as optimization techniques to reduce the computations necessary to accurately count the repetitions of such exercises as a guideline to future applications using wrist-worn devices.

II. Related Works

While exercise routines and detection are a popular field of research, ranging from quality to energy expenditure, several related works aim to detect particular exercises and count their repetitions. In [32], a mobile phone platform is used to aggregate data from two custom accelerometer devices operating at 100 Hz. These sensors, on the arm and the leg, as well as a heart rate sensor on the chest, show the ability to count repetitions of exercises and calories burned from the increased heart rate through the use of a six-dimensional Gaussian distribution for each movement class. They show an accurate system from 71% to 100% for 16 various activities using said system and counting the repetitions of an activity through a peak-detection method, which may or may not be a robust method when used across different populations. This work will attempt to approach a similar problem from that of [32] by using the classifiers to identify not the general motion patterns but instead use the classification to accurately count the motions detected. Further, this work will identify the key features in classifying free-weight and body-weight exercises.

In [33], a smartphone is attached to the users arm for unconstrained exercises (those capable of being performed in any environment) and placed on the weights of a machine in a constrained (gym) environment. They use a method based upon dynamic time warping to identify the activity, and possible repetitions, then count off of this information. In particular, they indicate that the DTW is too time-consuming to be performed real-time, and thus, pre-filter data by using peaks of similar height and a threshold window to only consider given acceleration windows. The similarity method appears robust in their results resulting in perfect precision and high recall (above 93% in all cases) in their unconstrained environment. This paper, similarly, deals with an unconstrained environment and movements monitored on the wrist (5 in their case) and attempts to find an even less-computationally intensive procedure as DTW to classify its movements. Further, no gyroscope exists in their platform, but will be analyzed here similarly to [25]. This work attempts to classify each individual movement repetition instead of the entire data set, similar to what was achieved in [33].

In [23], two different classifiers are used in determining a workout exercise then counting the repetitions of said exercise. Using a single accelerometer on the back of the hand and one on the hip, they are able to accurately identify most work out routines with most exercise accuracies in the $90-100\%$ range using a Naive Bayes classifier or a Hidden Markov Model. Further, after identifying the movement, they use either a peak detection algorithm, or the state transitions of their HMM in order to count the actual repetitions and have an error rate under 10% for most exercises and under 20% for all. They test their system in two different cross-validation schemes, with user-specific models at different weights and in a leave-one-subject-out cross validation for robustness, showing similar results. This work attempts to build upon results of [23], by assuming first that the motion may be identified correctly, either through their method, or through knowledge of the workout routine to identify the activity, and use a classifier to improve upon the counting error rates. Further, they eliminate the gyroscope, stating cost issues. This work will use a gyroscope to identify accuracies so that, if cost is not a concern, results can be evaluated to determine if it is necessary to include all available sensors in a smartwatch platform. Finally, instead of using the entire data set to separate the differences between motion classes, this work looks at individual repetitions and attempts to classify each.

Fig. 1. Device worn on wrist for exercise with dumbbell

TABLE I. Movements selected and whether they use a dumbbell or not

Movement	Uses Dumbbell
Bicep Curls	yes
Crunches	no
Jumping Jacks	no
Push Ups	no
Shoulder Lateral Raises	yes

III. Methods

TABLE II. List of Calculated Features per Axis

Feature	Count
Amplitude	6
Median	6
Mean	6
Maximum	6
Minimum	6
Peak-to-Peak	6
Variance	6
St. Dev.	6
Root Mean Square (RMS)	6
Skewness	6
Derivative Mean	6
Derivative St. Dev.	6
Derivative Variance	6
Derivative RMS	6
Axis Correlations	15

Many workout routines start and end the same way with

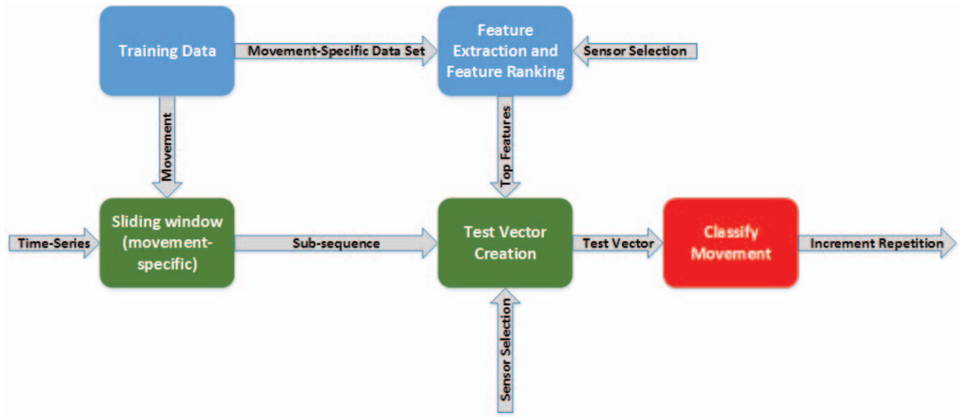


Fig. 2. Feature Ranking and Classification Method for Identifying Repetitions

a defined set of exercises. This work aims to leverage this information to build classification and repetition counting into a single algorithm to quantify these exercises. This effort will focus on building binary classifiers for each specific repetition of each movement, classifying individual actions instead of the repetitive pattern of the general exercise. Using a smartwatch platform that consists of an accelerometer and a gyroscope [29] operating at 50 Hz, shown heuristically to be +/- 2g and +/- 200 degrees per second respectively, worn on the left wrist. Fig. 1 shows a user wearing the watch during a workout routine while wearing the watch. A classification algorithm is run to determine accurately the workout routine movements listed in Table I. When looking at the watch face (e.g., when looking at the watch to read the current time), the x-axis points forward to the top of the watch, the y-axis points to the right, and the z-axis faces straight up from the face of the watch. While some of the routines could be done with or without dumbbells, the exercises selected were suggested by Focus [31] as a core set of initial exercises, using a 10 pound dumbbell. The end goal of the training and testing sets is to then develop an algorithm that is capable of recognizing these activities while running on the same watch.

A. Data Set

Data was collected on 12 participants, male and female, ranging from age 23 to age 38 and thus a lower weight was selected to account for all users. Once the data was collected, each was annotated manually for the start and end of each repetition. Each user was asked to perform each given activity ten times. The data was collected with the watch on the left hand (although the algorithms can be adapted to either hand). Finally the users were also asked to simply perform no movement, random movement, and various arm positions in order to develop a no-movement class for an application. Random windows from this signal were selected for the training sets.

B. Training

1) Preprocessing: Once the data was collected, the appropriate window size for each movement needs to be determined. This was done by calculating the average window size. For each move $m \in M$, where M is the set of all movements, each person $p \in P$, where P is the set of all people, has a move size determined by:

$$w_{m_p} = \frac{1}{n}\sum_{i=1}^{n} s_i \tag{1}$$

where s_i is each individual move sample defined by the annotated start and end points. This average is then calculated and the movement average is determined by:

$$w_m = \frac{1}{|P|}\sum_{j=1}^{|P|} w_{m_p} \tag{2}$$

This average window size is then used to alter the end point of each movement annotation to give the general movement size of each exercise. Since the workout routine is known, the context can be leveraged to build binary classifiers. As a user goes through a weight training program, the sets are defined. As a result, every model can be built to identify if a given movement window is a push up or is not a push up. Using this context information a series of features are extracted for each move.

2) Feature Extraction: The total list of features shown in Table II are extracted for each training sample. These features listed are calculated on each axis and the total number of the features is shown in the 2nd column. Thus, 6 features equates to the mean being calculated on each axis of gyroscope followed by each axis of the accelerometer. Once this is extracted for each sample for each person and each move, the number of features used in the testing must be reduced. Clearly the 99 features calculated would over-fit the data. Also, for computational performance it would be better to only need to calculate a subset of the features. Four testing configurations were created to evaluate the performance of the accelerometer, the gyroscope and the best axis. These configurations create 4 training sets per activity. The first is using features from only the accelerometer, the second only the gyroscope, the third is the combination of features from accelerometer and gyroscope, and finally the fourth is identifying the single best axis from the third set, and using only those features. Since there are significantly larger quantity of negative examples than positive for each movement, the negative samples are

TABLE III. AVERAGE ACCURACY AND AREA UNDER THE (ROC) CURVE (AUC) OF EACH MOVEMENT IN CROSS-VALIDATION FOR EACH OF THE TEST SETTINGS

Movement	**Accel. Only** Acc, AUC	**Gyro. Only** Acc, AUC	**Accel+Gyro** Acc, AUC	**Best Single Axis** Acc, AUC
Bicep Curls	92%, 0.97	87%, 0.91	92%, 0.98	87%, 0.94
Crunches	98%, 0.99	86%, 0.92	98%, 0.99	93%, 0.99
Jumping Jacks	88%, 0.94	77%, 0.86	89%, 0.95	84%, 0.94
Push Ups	96%, 0.99	87%, 0.93	96%, 0.99	95%, 0.99
Shoulder Lateral Raises	88%, 0.95	90%, 0.94	90%, 0.96	90%, 0.94

TABLE IV. SELECTED FEATURES OF EACH BEST AXIS FOR EACH MOVEMENT

Movement	Axis	Features
Bicep Curls	a_z	minimum, median, mean, amplitude, maximum, root mean square
Crunches	a_y	median, mean, maximum, minimum, variance, root mean square
Jumping Jacks	a_x	minimum, root mean square, median, st. dev., mean, peak-to-peak dist.
Push Ups	a_z	median, mean, maximum, minimum, variance, root mean square
Shoulder Lateral Raises	g_y	root mean square, st. dev., minimum, peak-to-peak dist., variance, maximum

randomly removed to balance the training set size. Finally, then, the training and testing sets are created to be used with feature ranking and selection, then classification. The training and classification flow is shown in Fig. 2. The features, once extracted, are normalized. In order to make the correlation features more informative, the signals are normalized for the extraction of those features.

C. Testing

1) Cross-Validation: In order to test the model, the Weka [34] platform was used in order to feature rank, feature select, and cross-validate the results. The top six features were used in order to ensure that the accuracy information will not be as a result of over-fitting the training data. Six features were chosen as the minimal subset that achieves what was deemed acceptable performance. Using a correlation ranker, the top six features were used in each set and the results were past through a 10-fold cross validation to test robustness to variability. Several classification schemes can be tested in Weka, and in this case random forests, decision trees, SVM, and Naive Bayes classifiers were compared.

2) Testing in a Real Setting: For the test data, a sliding window of points, equal to the average window size calculated for each movement is chosen. Given a time series signal $T = t_0, t_1, ...$ where each index is a sample at 50Hz, a subsequence t_i is then selected as follows:

$$t_i = (T_i, T_{i+w}) \tag{3}$$

where w is the indicated move window as designated from the training set, assuming enough points remain in the time-series. This subsequence is then used to extract features, resulting in a feature test-vector, as in:

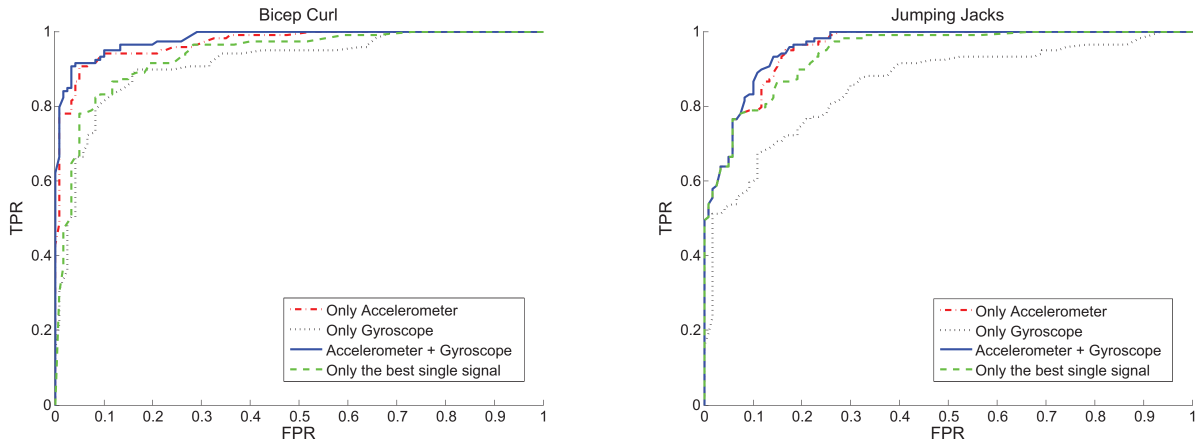
$$F(t_i) = \{f_i | 0 < i \leq MAX_f, each f_i \in F_s\} \tag{4}$$

where MAX_f indicates the largest number of features collected, then each f_i is from the sorted order of features as ranked by the feature ranking algorithm, denoted F_s, such that f_0 is the highest ranked feature, f_1 the next and so on. A sliding window is then shifted over the testing data testing each activity. Every time an activity is classified as an individual sample, then it is counted as a repetition, the window cleared and the time-series sub-sequence jumps past this point and continues (in order to avoid re-classifying the same repetition when the window is slid only one point). Otherwise, the overlap for the sliding window is every point. For every yes classification result a counter is incremented to count repetitions. Thus, high classification accuracy of individual movement repetitions will equate with low error rate in counting.

IV. RESULTS

3) Accuracy and AUC: In order to evaluate the results the correlation tool to rank the features in the Weka environment was used, followed by a random forest classifier due to its highest classification accuracy. The top six features were picked for each movement. As seen in Table III, all testing configurations provide strong results. This table shows the classification accuracy as well as the area under the curve (AUC) of the Receiver Operating Characteristics (ROC) of each movement in each of the four test settings. The gyroscope seems to add some strength in a few of the movements and thus, for a system that wants accurate counting, is necessary in order to have the highest classification accuracy. The jumping jacks, in particular, have curves that occasionally saturate the 2g accelerometer, thus, most likely exhibits the lowest classification accuracy. Thus, by using context information to create binary classifiers, a high classification rate for each example movement is found, thus indicating a low error rate on counting for each of these movements in a trainer application.

4) Selecting the Best Axis: In order to reduce the computational load, a trade off between the best single axis and the full data set can be compared. Two examples of movements and their ROC curves are shown in Fig. 3a and Fig. 3b, showing a combination of accelerometer only, gyroscope only, accelerometer and gyroscope, as well as the best selected individual axis. Note, in these examples, a combination of accelerometer and gyroscope provide the strongest result. The best individual axis and their features for each movement are listed in Table IV, listed in order of strongest first. While the features are similar, a binary classifier allows for the selection of different features for each movement. Thus, contextual information can optimize a system to use the ideal features and reduce computational load. Further, by only attributing a specific axis, less data can be computed or transmitted, and further, uniaxial sensors can be chosen in place of triaxial if so desired. For certain movements, this procedure yields accurate results that allow for the reduction of the computational load. Fig. 4 shows one such movement where the accelerometer



(a) ROC Curves for Bicep Curls

(b) ROC Curves for Jumping Jacks

Fig. 3. ROC Curves for (a) Bicep Curls and (b) Jumping Jacks showing Accelerometer Only, Gyroscope Only, Accelerometer + Gyroscope, and Best Axis Classification Methods

only and the accelerometer plus gyroscope configurations use the same best features as the single individual axis of the accelerometer. This allows for the same ROC curve and same AUC as a result.

V. Future Work

The system shown here provides ample opportunity to further test and evaluate workout routines. Further feature extraction techniques can be used to determine stronger best-axis information to reduce the computational load, or to ensure a minimal subset of considered sensor data, as well as applying contextual information or calibration steps to create a user-centric platform. Once this is developed, further detailed motions can be analyzed in a similar manner, such as those needed for rehabilitative exercises. As the workout routines are increased this may be more important and so more exercises should be considered, as well as the real-time responsiveness and user-experience considered. Further, the algorithm itself should be re-implemented on to the watch and a user-experience trial performed to see if users appreciate the accuracy and speed in which the repetitions are counted, and the actual delay values calculated. Such systems should also be tested for robustness in terms of varying weights, in particular considering the trade-offs between movement-specific models and weight-class specific models and how this contextual information may be passed into the system. Finally, the addition and fusion of other sensors, such as heart rate, may provide valuable differentiating characteristics and should be further investigated in the realm of exercise-sport applications.

VI. Conclusion

Smartwatches show a new realm of activity sensors for personal health monitoring. Some have accelerometers while others also include a gyroscope. As wrist-worn wearables increase in popularity, the investigation of the sensor platforms needed and computational considerations for the response time of such counting algorithms must be considered, ranging in applications from general monitoring to real-time exercise repetition. This work presents the feature extraction and selection necessary to not only identify but at the same time count repetitions of free-weight and body-weight exercises by leveraging the context of the workout routine in order to develop strong classifiers. While the accelerometer alone is fairly consistent in classifying the individual repetitions of the motions, the gyroscope does improve the classification accuracy by having an average area under the curve of .974, up from .968 using only the accelerometer. Adding a gyroscope with the accelerometer increased the average AUC from .968 to .974, increasing the accuracy of specific movements as much as 2%. The best single axis method achieves an AUC of .95 showing that a reduced method can provide the necessary accuracy to accurately count repetitions in a lower computational power setting.

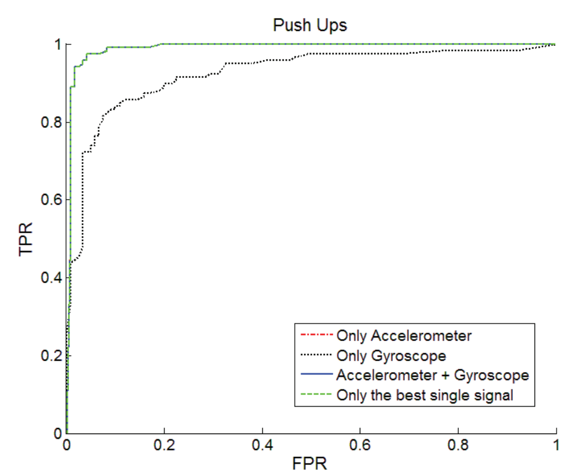


Fig. 4. ROC Curves for push ups. Note that three configurations share the same curve

ACKNOWLEDGMENT

The authors would like to thank the team at Focus, Cavan Canavan and Grant Hughes, for their input on ideal initial exercises and the workout routine.

REFERENCES

[1] S. Wu and G. A.P., "Projection of chronic illness prevalence and cost inflation," 2000.

[2] T. Bodenheimer, E. H. Wagner, and K. Grumbach, "Improving primary care for patients with chronic illness," *JAMA: the journal of the American Medical Association*, vol. 288, no. 15, pp. 1909–1914, 2002.

[3] F. W. Booth, C. K. Roberts, and M. J. Laye, "Lack of exercise is a major cause of chronic diseases," *Comprehensive Physiology*, 2012.

[4] P. A. Heidenreich, J. G. Trogdon, O. A. Khavjou, J. Butler, K. Dracup, M. D. Ezekowitz, E. A. Finkelstein, Y. Hong, S. C. Johnston, A. Khera *et al.*, "Forecasting the future of cardiovascular disease in the united states a policy statement from the american heart association," *Circulation*, vol. 123, no. 8, pp. 933–944, 2011.

[5] A. D. Association *et al.*, "Economic costs of diabetes in the u.s. in 2012. diabetes care 2013; 36: 1033–1046," *Diabetes Care*, vol. 36, no. 6, p. 1797, 2013.

[6] F. Pitta, T. Troosters, M. A. Spruit, V. S. Probst, M. Decramer, and R. Gosselink, "Characteristics of physical activities in daily life in chronic obstructive pulmonary disease," *American journal of respiratory and critical care medicine*, vol. 171, no. 9, pp. 972–977, 2005.

[7] M. Buman, J. Kurka, E. Winkler, P. Gardiner, E. Hekler, G. Healy, N. Owen, C. Baldwin, and B. Ainsworth, "Estimated replacement effects of accelerometer-derived physical activity and self-reported sleep duration on chronic disease biomarkers," *Journal of Science and Medicine in Sport*, vol. 15, p. S76, 2012.

[8] E. Bacchi, C. Negri, G. Targher, N. Faccioli, M. Lanza, G. Zoppini, E. Zanolin, F. Schena, E. Bonora, and P. Moghetti, "Both resistance training and aerobic training reduce hepatic fat content in type 2 diabetic subjects with nonalcoholic fatty liver disease (the raed2 randomized trial)," *Hepatology*, vol. 58, no. 4, pp. 1287–1295, 2013.

[9] M. L. Pollock, B. A. Franklin, G. J. Balady, B. L. Chaitman, J. L. Fleg, B. Fletcher, M. Limacher, I. L. Piña, R. A. Stein, M. Williams *et al.*, "Resistance exercise in individuals with and without cardiovascular disease benefits, rationale, safety, and prescription an advisory from the committee on exercise, rehabilitation, and prevention, council on clinical cardiology, american heart association," *Circulation*, vol. 101, no. 7, pp. 828–833, 2000.

[10] S. Chen, J. Lach, O. Amft, M. Altini, and J. Penders, "Unsupervised activity clustering to estimate energy expenditure with a single body sensor," in *Body Sensor Networks (BSN), 2013 IEEE International Conference on*, May 2013, pp. 1–6.

[11] F. Dadashi, A. Arami, F. Crettenand, G. P. Millet, J. Komar, L. Seifert, and K. Aminian, "A hidden markov model of the breaststroke swimming temporal phases using wearable inertial measurement units," in *Body Sensor Networks (BSN), 2013 IEEE International Conference on*, May 2013, pp. 1–6.

[12] G. Plasqui, A. Bonomi, and K. Westerterp, "Daily physical activity assessment with accelerometers: new insights and validation studies," *Obesity Reviews*, 2013.

[13] S. Liu, R. Gao, and P. Freedson, "Computational methods for estimating energy expenditure in human physical activities." *Medicine and science in sports and exercise*, 2012.

[14] C. V. Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 3, pp. 136–147, 1997.

[15] S. E. Crouter, K. G. Clowers, and D. R. Bassett, "A novel method for using accelerometer data to predict energy expenditure," *Journal of applied physiology*, vol. 100, no. 4, pp. 1324–1331, 2006.

[16] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys (CSUR)*, vol. 46, no. 3, p. 33, 2014.

[17] K. Lyden, S. L. Kozey, J. W. Staudenmeyer, and P. S. Freedson, "A comprehensive evaluation of commonly used accelerometer energy expenditure and met prediction equations," *European journal of applied physiology*, vol. 111, no. 2, pp. 187–201, 2011.

[18] H. Kalantarian, S. I. Lee, A. Mishra, H. Ghasemzadeh, and M. Sarrafzadeh, "Multimodal energy expenditure calculation for pervasive health: A data fusion model using wearable sensors," in *Proceedings of IEEE PerCom Workshop on Smart Enviornments and Ambient Intelligence*. ACM, 2013.

[19] A. J. Daley, "Can exergaming contribute to improving physical activity levels and health outcomes in children?" *Pediatrics*, vol. 124, no. 2, pp. 763–771, 2009.

[20] B. Mortazavi, S. Nyamathi, S. Lee, T. Wilkerson, H. Ghasemzadeh, and M. Sarrafzadeh, "Near-realistic mobile exergames with wireless wearable sensors," *Biomedical and Health Informatics, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2013.

[21] W. Peng, J.-H. Lin, and J. Crouse, "Is playing exergames really exercising? a meta-analysis of energy expenditure in active video games," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 11, pp. 681–688, 2011.

[22] M. Kranz, A. Möller, N. Hammerla, S. Diewald, T. Plötz, P. Olivier, and L. Roalter, "The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices," *Pervasive and Mobile Computing*, vol. 9, no. 2, pp. 203–215, 2013.

[23] K.-H. Chang, M. Y. Chen, and J. Canny, "Tracking free-weight exercises," in *UbiComp 2007: Ubiquitous Computing*. Springer, 2007, pp. 19–37.

[24] K. S. Choi, Y. S. Joo, and S.-K. Kim, "Automatic exercise counter for outdoor exercise equipment," in *Consumer Electronics (ICCE), 2013 IEEE International Conference on*. IEEE, 2013, pp. 436–437.

[25] J. Parkka, M. Ermes, K. Antila, M. van Gils, A. Manttari, and H. Nieminen, "Estimating intensity of physical activity: a comparison of wearable accelerometer and gyro sensors and 3 sensor locations," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 1511–1514.

[26] A. Moller, L. Roalter, S. Diewald, J. Scherr, M. Kranz, N. Hammerla, P. Olivier, and T. Plotz, "Gymskill: A personal trainer for physical exercises," in *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*. IEEE, 2012, pp. 213–220.

[27] E. Velloso, A. Bulling, H. Gellersen, W. Ugulino, and H. Fuks, "Qualitative activity recognition of weight lifting exercises," in *Proceedings of the 4th Augmented Human International Conference*. ACM, 2013, pp. 116–123.

[28] C. Li, M. Fei, H. Hu, and Z. Qi, "Free weight exercises recognition based on dynamic time warping of acceleration data," in *Intelligent Computing for Sustainable Energy and Environment*. Springer, 2013, pp. 178–185.

[29] Samsung, "Samsung galaxy gear." [Online]. Available: http://www.samsung.com/levant/consumer/mobile-phones/mobile-phones/galaxy-gear/SM-V7000ZWAAFR-spec

[30] Sony, "Sony smartwatch 2." [Online]. Available: http://www.sonymobile.com/global-en/products/accessories/smartwatch-2-sw2/specifications/

[31] Focus, "Focus digital trainr." [Online]. Available: http://www.focustrainr.com/

[32] C. Seeger, A. Buchmann, and K. Van Laerhoven, "myhealthassistant: a phone-based body sensor network that captures the wearer's exercises throughout the day," in *Proceedings of the 6th International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011, pp. 1–7.

[33] I. Pernek, K. A. Hummel, and P. Kokol, "Exercise repetition detection for resistance training based on smartphones," *Personal and ubiquitous computing*, vol. 17, no. 4, pp. 771–782, 2013.

[34] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.