FitCoach: Virtual Fitness Coach Empowered by Wearable Mobile Devices

Xiaonan Guo, Jian Liu, Yingying Chen
Department of Electrical and Computer Engineering
Stevens Institute of Technology, Hoboken, NJ, USA
Emails: {xguo6, jliu28, yingying.chen}@stevens.edu

***Abstract*—Acknowledging the powerful sensors on wearables and smartphones enabling various applications to improve users' life styles and qualities (e.g., sleep monitoring and running rhythm tracking), this paper takes one step forward developing FitCoach, a virtual fitness coach leveraging users' wearable mobile devices (including wrist-worn wearables and arm-mounted smartphones) to assess dynamic postures (movement patterns & positions) in workouts. FitCoach aims to help the user to achieve effective workout and prevent injury by dynamically depicting the short-term and long-term picture of a user's workout based on various sensors in wearable mobile devices. In particular, FitCoach recognizes different types of exercises and interprets fine-grained fitness data (i.e., motion strength and speed) to an easy-to-understand exercise review score, which provides a comprehensive workout performance evaluation and recommendation. FitCoach has the ability to align the sensor readings from wearable devices to the human coordinate system, ensuring the accuracy and robustness of the system. Extensive experiments with over 5000 repetitions of 12 types of exercises involve 12 participants doing both anaerobic and aerobic exercises in indoors as well as outdoors. Our results demonstrate that FitCoach can provide meaningful review and recommendations to users by accurately measure their workout performance and achieve 93% accuracy for workout analysis.**

I. Introduction

The proliferation of *wearable mobile devices* (e.g., smartwatches, wrist-worn fitness bands, and smartphones mounted on arms) has already shown its potential on improving our life styles through a great number of applications in smart healthcare, smart home, and smart cities. An important use case of wearable mobile devices is providing guidelines to improve people's daily activities, for example, tracking walking steps [18], monitoring sleep qualities [13], and estimating daily caloric intake [14]. In this work, we take one step forward by answering the question: Whether such wearable mobile devices become powerful enough leveraging fine-grained sensing information to perform systematic comprehensive fitness assistance and prevent injuries.

Traditionally, fitness monitoring is performed by analyzing the workout captured by video tapes [6] or specialized sensors [7], [8]. Chang *et al.* [7] track free-weight exercises by incorporating an accelerometer into a workout glove. Cheng *et al.* [8] develop a technique that can recognize human activities by attaching a sensor on users' hips. In recent years, smartphone apps, fitness trackers and dedicated devices, such as Sworkit [5], Fitbit [1], Garmin watch [2] and Gym watch [3], show the initial success of fitness monitoring. They can perform step counts and log exercises based on users' manual inputs. Additionally, people need to purchase dedicated sensors and wear them during exercises. Hao *et al* [11] present a system using smartphone and its external microphone that detects running rhythm and improves exercise efficiency for runners, yet the question whether or not mobile devices can automatically distinguish different types of exercises and provide fine-grained performance recommendation related to exercises remains open.

Toward this end, we take one step forward to search for an integrated mobile solution that can perform systematic fitness monitoring and performance review. We propose *FitCoach* leveraging wearable mobile devices to achieve the following two main aspects: **(i) Fine-grained Fitness Data Interpretation.** Recording the sensor readings on wearable mobile devices (e.g., smartwatch or smartphone) during workout to explore their capability of deriving fine-grained exercise information including exercise types, the number of set and the number of repetitions (reps) per set. The derived quantitative data can be further analyzed for inferring meaningful information. For example, higher level information can be obtained including calories burn, body fat, body mass index, etc. **(ii) Smart Exercise Guidance.** Furthermore, the derived fitness data is of great importance to assist the users to maintain proper exercise postures and avoid injuries. To build muscles and gain a healthier body, it is widely recognized that people should perform their workout properly and effectively. FitCoach aims to not only regulate the workouts by following the Frequency, Intensity, Time and Type (FITT) principle [16], but also provide detailed guidelines to review the user's posture through workout and provide recommendation in keeping correct exercise form (e.g., in terms of speed of exercise execution and strength).

In particular, FitCoach exploits *Short Time Energy* (STE) to derive fine-grained fitness data (i.e., strength and speed of body movements) in exercises and recognizes different types of exercises automatically by using embedded sensors (e.g., accelerometer and gyroscope) on wearable mobile devices. Rooted in the understanding of body movements in exercises, FitCoach develops a novel metric for evaluating the quality of each user's exercises, *exercise form score*. This exercise form score reflects the difference of strength and speed of body movements between each repetition of an exercise based on a reference profile. The reference profile could be either obtained from the user's own sensor data or built from other people's data (e.g., training coaches or members from the same fitness club) through crowdsourcing platforms (e.g., fitness club's facebook, WhatsApp or WeChat).

The contributions of our work are summarized as follows:

- Assessing dynamic postures (movement patterns & positions) automatically during workout including anaerobic as well as aerobic exercises.
- Achieving fine-grained exercise recognition (including exercise types, the number of sets and repetitions) without user involvement.
- Calculating exercise form score and providing performance review to assist high-quality workout and prevent injuries for both short term and in the long run.
- Aligning sensing data into the human coordinate system to ensure high recognition accuracy and achieve system robustness even when the real-time data possess the different device facing direction or exercise direction comparing to the reference profiles.
- Demonstrating the system performance involving 12 people using both smartwatches and mobile phones in armbands during both gym and outdoor workouts with high accuracy over 90% for workout analysis.

II. Related Work

Recent studies show that life experience can be improved through implementing various types of techniques using sensors and wireless technologies including activity recognition [8], [15], [17], [20], [22], [23] and physical exercises monitoring [1], [7], [10], [11], [15].

There has been active work for activity recognition, including daily activities [8], [15], [22] and healthcare related activities such as eating [20] and smoking [17]. Vlasic *et al.* [22] develop a full body motion capture system by using multiple sensors attached on a human body. Cheng *et al.* [8] develop a technique that can recognize activities without training by placing a sensor on users' hips. These studies show that either external sensors or sensors embedded in wearables have the capability to accurately recognize human daily activities. Furthermore, video-based technologies can capture and recognize human hand motion [19] but require line-of-sight.

Another aspect of related studies focus on automatically monitoring physical exercises. There are mobile Apps [5], wristband [1] and solutions based on mobile devices with sensors [7], [10], [11], [15]. Chang *et al.* [7] propose to track free weight exercises by incorporating an accelerometer into a workout glove. In addition, Ding *et al.* [10] propose to recognize free-weight activities by attaching passive RFID tags on the dumbbells. Along this line, Hao *et al.* [11] propose to monitor the running rhythm by measuring breathing and strides with headsets and smartphones. These techniques rely on additional sensors or specific hardware. Most importantly, whether a workout feedback and guidance can be further provided to improve exercise performance still an open question.

The commercial products also exhibit the trend to automate the fitness monitoring, such as Garmin watch [2] and Gym watch [3]. However, Garmin watch requires explicit inputs from users, including the type of workout and the start/stop time. Gym watch requires people to purchase dedicated sensors and wear them during exercises. Along this trend, FitCoach takes one step forward by utilizing the existing wearable devices (e.g., wrist-worn smartwatches or arm-mounted smartphones) to automatically provide fine-grained tracking of workout and offer exercise review and guidance to improve fitness experience.

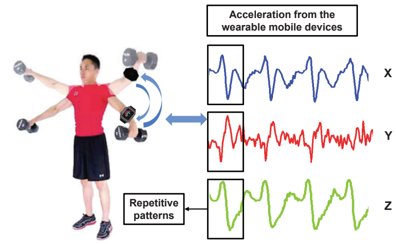


Fig. 1. Movement in exercises can be revealed by repetitive patterns from sensor readings of wearable mobile devices.

III. Design of FitCoach

A. Challenges and Practical Issues

Exercise Form Correction Using Single Wearable Mobile Device. It is necessary for the system to understand the performance of a exercise through the body movements, which is a challenging task to cope with by using a single wearable mobile device. This is because commercial mobile devices usually have limited low-power sensing modalities (i.e., accelerometer, gyroscope and magnetometer). Therefore, the system needs to be designed in such a way that can provide exercise form corrections based on the dynamics of sensor data resulted from the partial knowledge of the exercises.

Robust Fine-grained Exercise Differentiation. It is also challenging to utilize sensors in wearable mobile devices to correctly distinguish different types of exercises, since sensor readings collected from the wearable mobile devices are extremely noisy due to the dynamic nature of exercises. Thus, it is important to devise a robust exercise classifier that can eliminate the impact of noisy sensor data and capture the fine-grained differences between different types of exercises.

Automated Wearing Orientation Alignment. During exercises, wearable mobile devices may change its facing from the original direction from time to time. Such orientation changes result in unstable projection of user's body movements in the mobile device's coordinate system, and makes it hard for the system to determine the pattern of body movements. Therefore, a light-weight alignment algorithm is needed to transform the sensor data to that in a stable orientation to facilitate accurate exercise recognition.

B. System Overview

The main goal of FitCoach is to examine the users' dynamics (i.e., body movement patterns & intensities) in workouts and provide detailed workout statistics to assist users to achieve effective workouts and prevent injuries.

Given that these wearable mobile devices are worn on the human body of either wrist or upper arm, they become

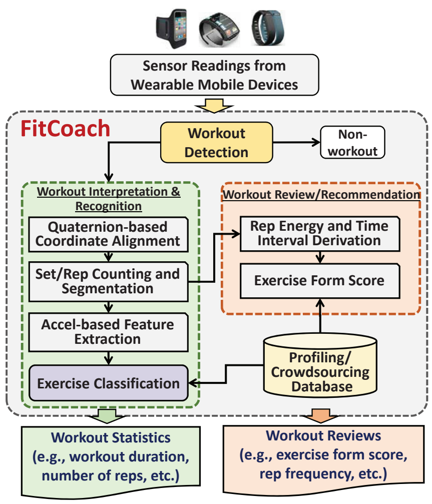


Fig. 2. FitCoach framework.

desirable interfaces to sense exercise movements to provide detailed workout statistics/analysis. As illustrated in Figure 1, the repetitive pattern of body movements in exercises can be well captured by using the inertial sensors of the wearable mobile device (i.e., a smartwatch). FitCoach can automatically extract fine-grained fitness information (e.g., basic statistics, motion energy and performing period) without users' cooperation and provide illustrative feedback to users, which can also be exploited to enforce the Frequency, Intensity, Time, Type (FITT) principle of training [16].

As illustrated in Figure 2, FitCoach takes as input time-series of sensor readings from accelerometer and gyroscope as well as quaternion, all of which are readily available in off-the-shelf wearable mobile devices. We first perform *Workout Detection* to filter out the sensor readings that don't contain workout activities based on the presence of periodicity pattern in workout activity. The sensor readings that are found to contain workout activities will be served to two tasks, *Workout Interpretation & Recognition* and *Workout Review/Recommendation*. The Workout Recognition performs quantitative analysis to the sensor readings and identify different types of workouts based on the acceleration features that can capture unique repetitive patterns of different exercises. The Workout Review/Recommendation examines the characteristics of each rep (i.e., energy and time intervals) and provides the novel exercise form scores as feedback to users for performance evaluation.

Particularly, the Workout Recognition consists of four major components: *Quaternion-based Coordinate Alignment*, *Set/Rep Counting and Segmentation*, *Accel-based Feature Extraction*, and *Exercise Classification*. The Quaternion-based Coordinate Alignment tackles the issue of dynamic orientation in workouts, and automatically rotates sensor readings to a fixed coordinate system. The Set/Rep Counting counts the number of sets during the workout and the number of reps in each set based on the magnitude of the repetitive signals resulted from workouts. The sensor readings are further divided into small segments corresponding to the detected reps. In each segment,

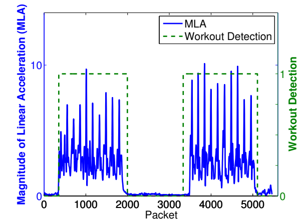


Fig. 3. Workout detection based on a 5-second sliding window (output 1 if the number of repetitive patterns is larger than 3 within the window, otherwise output 0).

the Accel-based Feature Extraction derives statistics features that capture each repetitive moving patterns of exercises from three-axis acceleration readings. After Workout Interpretation, the system performs *Exercise Classification*, which utilizes a profile based algorithm to determine the types of exercises by comparing the extracted features with those of pre-collected profiles in the *Profiling/Crowdsourcing Database*.

In addition, the Workout Review/Recommendation aims to provide systematic fitness monitoring and performance review as feedback to users, which would assist the users to maintain proper exercise gestures and avoid injuries. FitCoach takes the segments of sensor readings identified in the Set/Rep Counting and Segmentation as inputs, and performs the *Rep Energy and Time Interval Derivation* to estimate the characteristics of body movements in exercises (i.e., strength and frequency of the repetitive motions). The estimated characteristics are further utilized by the *Exercise Form Score Calculation* to calculate the exercise form score for each rep, which is a novel metric that allows the users to easily understand their performance in the exercises.

IV. Workout Interpretation & Recognition

A. Workout Detection

A key observation is that most regular exercises involve repetitive arm movements. For example, jogging and walking involve periodic arm swing, and weight lifting involves periodic pushing-ups. Such repetitive arm movements result in regularly changing values in sensor readings. In addition, the repetitive patterns from exercises tend to be last for a long time period simply because people normally adopt a set-and-rep scheme in exercise to maximize the effectiveness. Compared to regular exercises, non-workout activities usually don't have such long-term repetitive pattern. Therefore, we propose to detect workout based on determining whether there are long-term repetitive patterns in the sensor readings.

Towards this end, we adopt an autocorrelation-based approach to examine the accelerations resulted from exercise motions. The autocorrelation approach is a common technique used for detecting repetitive patterns in a time series. In particular, we first apply a moving time window with the length of w to the time series of accelerometer readings. For each time window, we use the Magnitude of Linear Acceleration (MLA) to estimate the linear acceleration (i.e., acceleration without gravitational acceleration) of exercise motions. The MLA based on accelerometer readings can be derived by the following equation:

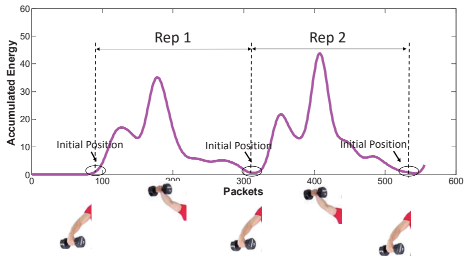


Fig. 4. Illustration of the relationship between the arm movements in a repetition and the unique pattern of accumulated energy captured by a wearable mobile device (i.e., a smartphone in an armband).

$$MLA(i) = \sqrt{(a(i)_x)^2 + (a(i)_y)^2 + (a(i)_z)^2} - g, \quad (1)$$

where $a(i)_x, a(i)_y$ and $a(i)_z$ are the acceleration of the ith sample on the x, y and z axis of the mobile device respectively and g is the acceleration of gravity. Note that, the MLA in Equation 1 equals to zero when there is no motion.

Then we calculate the autocorrelation of the time series of MLA, and use a typical peak finding algorithm [13] to find the number of peaks in the autocorrelation, which is denoted as N_p. The number of detected repetitive patterns thus can be derived with $N_r = (N_p - 1)/2$, due to the symmetric nature of the autocorrelation. Finally, to accommodate the noisy accelerometer readings, we use a threshold-based method to confirm the detected repetitive patterns are resulted from workouts. The workout detection results for each window can be derived by:

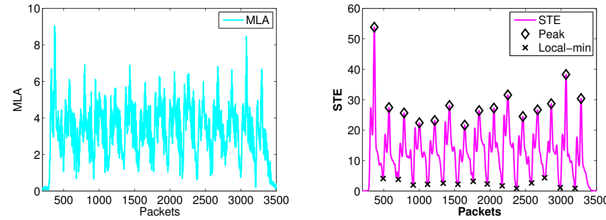
$$D_w = \begin{cases} 1, & N_r > \nu \\ 0, & otherwise, \end{cases} \quad (2)$$

where D_w is a boolean value depicts whether the given sensor readings within a window belong to workout or not. D_w outputs 1 when N_r is bigger than a threshold value ν. Figure 3 shows an example of our workout detection results with $w = 5$s and $\nu = 3$, which demonstrates that our system can accurately detect the windows containing workouts.

B. Set/Rep Segmentation

After the Workout Detection, FitCoach integrates the windows that are continuously labeled as workouts into a segment. The time between any two segments are identified as the rest interval, which will be provided as a part of the exercise review. However, in order to provide fine-grained exercise performance information, FitCoach needs to look into the data in each set and analyzes the data based on a finer-grained concept, *repetition/rep*.

We devise a motion-energy-oriented approach to accurately estimate starting and ending time point of each repetition of the same exercise motion within a set. The intuition behind the approach is that each repetition usually consists of a series of arm movements that result in a unique pattern in terms of the accumulated motion energy: 1) the accumulated energy starts to increase sharply from zero when the arm moves from an initial position to an ending position; 2) the accumulated energy drops a little when the arm pauses at the ending position for a very short while; 3) the accumulated energy starts to increase sharply again when the arm moves back from the ending position to the initial position; and 4) finally the accumulated energy drops sharply when the hand stops at the initial position for some rest. We found that this unique pattern of accumulated motion energy can be captured by the wearable mobile device through the Short Time Energy of MLA. Figure 4 illustrates the relationship of the unique pattern in the accumulated energy and the arm movements in each repetition.



(a) Magnitude of Linear Acceleration (MLA)

(b) Local minimum identified in Short Time Energy (STE) of MLA

Fig. 5. Example of rep segmentation for 10 repetitions of dumbbell raising exercise.

Particularly, we adopt the Short Time Energy (STE) [9] to capture the unique energy pattern in the time series of MLA. The basic idea of this step is to accumulate the energy of the MLA in short sliding windows. After obtaining STE of MLA, FitCoach applies the same peak finding algorithm used in Section IV-A to detect the peaks in STE. Then the system finds the local minimum point between two peaks as the ending point of each repetition, and the data between two detected ending points are defined as a segment of repetition. Figure 5 shows an example of determining the repetition segments based on the local minimum points that are detected in STE of MLA from a wearable mobile device (i.e., a smartwatch) when the user conducts 3 sets of dumbbell rasing with 10 repetitions per set. The results indicate that the motion-energy-based approach can accurately separate the data for each repetition.

C. Accel-based Feature Extraction & Workout Classification

After repetition segmentation, FitCoach aims to identify the workout type for each set. The basic idea is to build a database with the profiles for different types of workouts before the workout classification, then we use a profile-based approach to determine the workout type for each rep segment in the set, and further to infer the workout type of the entire set.

Accel-based Feature Extraction. In order to distinguish different types of workouts, we need to find the features that can capture the unique characteristics of each type of workouts. Based on our extensive feature selection studies, we finally determine nine statistical acceleration-based features that are most useful to distinguish different types of workouts, namely *skewness, kurtosis, standard deviation, variance, most frequently appear in the array, median, range, trimmean and mean*. To extract features without worrying about the variation of the mobile device's facing orientation, we first perform the earth-reference alignment to rotate all acceleration data to the earth coordinate system. The details of the earth-reference alignment are provided in Section VI-A. After the world-reference alignment, FitCoach extracts the nine acceleration-based features from the already aligned three-axis accelerations in each rep segment to describe the body movements. In total, we extract 27 features (i.e., nine features per axis) for each rep segment.

Light-weight Classifier. FitCoach utilizes a light-weight machine learning based approach to identify different types of workouts based on the acceleration-based features extracted from each rep segment. It is light-weight because the system only needs to determine the workout type for the first few rep segments within a set, and the workout type of the entire set of repetitions is identified as the majority decision based on the classification results from the first few rep segments. Specifically, we adopt a Support Vector Machine (SVM) classifier [21] with radial basis function kernel. The classifier is trained by the pre-collected profiles of different types of workouts, which is described in Section VI-C. We note that we utilize the classification results of the first five reps to determine the workout type of the entire set.

V. Workout Review and Recommendation

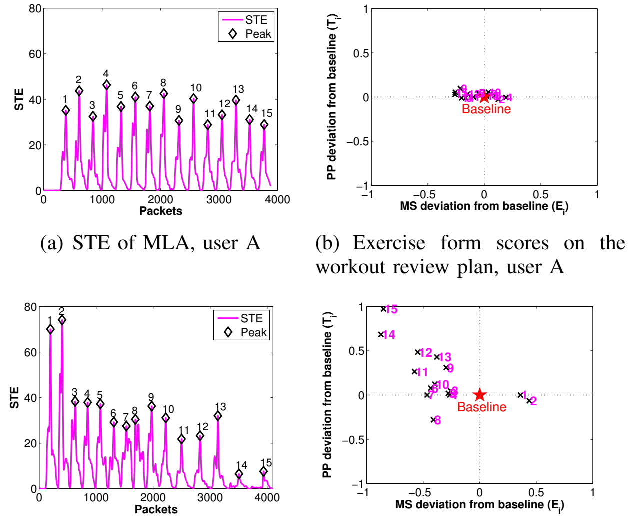
In this section, we first sketch the big picture of the workout review provided by FitCoach through summarizing the workout statistics, then discuss the details of our novel exercise form score and workout performance plane.

A. *Overview of Workout Review*

In order to achieve effective workouts and avoid injuries, users usually seek out personal fitness plans provided by fitness trainers or professionals. Such fitness plans often try to regulate the workouts by following the Frequency, Intensity, Time and Type (FITT) principle of training, which is a set of guidelines that instruct users to set up workout routines fitting their goals and fitness levels while maximizing the effects of exercises. However, most of users cannot afford a full-time personal trainer that can coach their workouts at any time. FitCoach fills the gap between users and the fitness plans based on FITT principle of training by providing fine-grained fitness information and intuitive feedback to users. Specifically, FitCoach is able to track the following basic workout statistics automatically including *exercise type, number of reps, number of sets, time between sets, time between sessions (training days/week)* to enforce the FITT principle of training. In addition, FitCoach further provides fine-grained feedback, which is the *exercise form score* in terms of motion energy and performance period for individual rep, to assist users in fine-tuning their exercises gestures.

B. *Exercise Form Score Design*

Besides providing basic workout statistics to the users, FitCoach aims to offer users a more intuitive way to understand their performance in exercises by comparing their exercise statistics to a baseline, which could be either generated based on the users' own data or based on the data from croudsourcing. Towards this end, we define a novel metric named *exercise form score*, which consists of two subscores that respectively evaluate a user's fine-grained performance of each rep in the exercise based on two important criteria as shown below:



(a) STE of MLA, user A

(b) Exercise form scores on the workout review plan, user A

(c) STE of MLA, user B

(d) Exercise form scores on the workout review plan, user B

Fig. 6. Comparison of the Short Time Energy (STE) of the Magnitude of Linear Acceleration (MLA) and the exercise form scores on the workout review plane between user A and user B.

Motion Strength (MS). A proper exercise form should maintain the motion strength at a certain level. For example, too much strength may indicate that the user spend more energy on each rep and if the weight is too heavy, it will increase the risk of injury while too little strength may indicate that the user spend too less energy to build muscle effectively. We intuitively utilize the energy level of each rep to describe the motion strength, which mean a set of reps with good performance should maintain a stable energy level. The energy level of each rep can be estimated by the maximum value in obtained STE of MLA.

Performing Period (PP). A proper exercise form should avoid too-fast or too-slow movements in order to effectively build muscles and prevent injuries. In this work, we utilize the time period of each rep to describe the performing period of each rep, which reflects how fast a user performs a repetition in exercises. Therefore, a set of reps with good performance should also have similar time periods. The time period of each rep can be directly obtained from the length of each rep segment after the segmentation described in Section IV-B. We note that the performing period provides more insights to users. For example, users can leverage such information for equipment weight adjustment (e.g., reduced speed of last few reps in a set indicates that the user may be training exhausted and need to decrease the weight or number of reps in next set).

Exercise Form Score. Based on these two criteria, FitCoach defines the *Exercise Form Score*, which consists of two subscores: *MS score* and *PP score*. The subscores depict how the testing rep deviates from the baseline in terms of the motion strength and performing period, respectively. We discuss the details about the baseline in the next subsection. Particularly, the MS score for the i^{th} rep is defined as:

$$E_i = \frac{A(i) - A^*}{A^*}, \quad i = 1, 2, 3, \ldots, n, \tag{3}$$

where $A(i)$ is the maximum STE of the MLA of the i^{th} rep, and A^* is the motion strength baseline. Similarly, the PP score for the i^{th} rep is defined as:

$$T_i = \frac{I(i) - I^*}{I^*}, \quad i = 1, 2, 3, \dots, n, \tag{4}$$

where I_i is the length of the i^{th} rep and I^* is the performing period baseline. The output exercise form score is a 2-tuple score that can be denoted as $< E_i, T_i >$.

C. Personal/Crowdsourcing Baseline

The exercise form score reflects the performance of the testing rep comparing to a baseline. We design two baselines that are suitable in different scenarios, namely *Personal Baseline* and *Crowdsouring Baseline*.

Personal Baseline. We observe that users usually can perform exercises with standard strength and frequency at the beginning of the workout, but the quality of the exercises decays with time due to fatigue. Based on this observation, a good candidate of the baseline for evaluating the performance of a user's workouts is the early portion of the user's own reps. In particular, we derive the personal baseline by averaging the motion strength and performing period of the first k reps of the first set in the user's sensor data. We empirically choose $k = 5$ in our work.

Crowdsourcing Baseline. The personal baseline is good for short-term exercise performance evaluation but could be bias to the user's own preference. For example, a user could feel tired at the beginning of the exercise and result in bad baseline for evaluating the entire exercise. To tackle this problem, we further propose the crowdsourcing baseline, which allows users to compare their performance with the baseline from exemplars (e.g., fitness coaches, bodybuilders, and amateur expertise) to achieve a long-term and more accurate exercise performance evaluation. The crowdsourcing approach is feasible because it is an increasing trend that people would like to share their fitness data in online social network to earn credits or build record, and more social platforms, such as WhatsApp and WeChat, start to provide the functionality allowing people to share their fitness data among friends.

D. Workout Review Plane

FitCoach further adopts an unique view angle of the exercise form score to allow users to track the performance or their each rep in a illustrative way. In particular, we define a *review plane* in which the x axis and y axis are the MS score and PP score, respectively. According to Equation 3 and 4, the Original represents the rep having the exactly same performance as the chosen baseline, and every exercise form score $< E_i, T_i >$ corresponding to the i^{th} rep can be mapped to a position in the the review plane. Apparently, the rep having its position closer to the Original has better performance, and the more reps close to the Original the better.

Figure 6 compares the workout reviews of two different users (i.e., User A an User B) in a set of lateral raising exercises (i.e., 15 reps in one set). Figure 6(a) and (c) respectively depict STE of MLA of two users' reps, which shows that User A has more stable energy levels and time lengths for each repetition than User B. Figure 6(b) and (d) respectively illustrate two users' exercise form scores based on their personal baselines in the review planes, which shows that the score points of User A are concentrated around the Original while the score points of User B are scattered around the second quadrant of the review plane. The observation indicates that User B have much higher motion strength and longer performing period comparing to the user's first few reps, and thus have worse performance than User A.

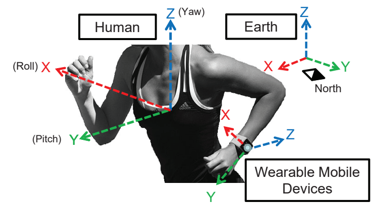


Fig. 7. Three coordinate systems.

VI. Implementation

A. Quaternion-based Coordinate Alignment

In workout monitoring scenarios, users wearing wearable mobile devices basically involve three different coordinate systems as illustrated in Figure 7, namely, *mobile device coordinate*, *earth coordinate*, and *human coordinate*. The sensor readings from a mobile device are defined in the device coordinate and thus result in non-fixed projection of the user's body movements defined in the human coordinate. In order to address this issue, FitCoach adopts a quaternion-based approache to dynamically convert sensor readings from the mobile device coordinate either to the human coordinate or to a coordinate system having the fixed mapping to the human coordinate.

1) Earth-reference Alignment: For exercise recognition in a gym, the orientation of wearable mobile devices may change due to rotation caused by arm movement. Therefore, our system needs to convert sensor readings from the mobile device coordinate to the earth coordinate first. Specifically, we convert the sensor readings from the mobile device coordinate to the earth coordinate by using the quaternion-based rotation $p_e = q_{me} p_m q_{me}^{-1}$, where p_m is the sensor reading vector (e.g., accelerations) in the mobile device coordinate, and q_{me} is the quaternion reading from the mobile device coordinate to the earth coordinate, which can be obtained from the device directly. q_{me}^{-1} is the conjugate quaternion of q_{me}. After conversion, the converted sensor readings p_e are in the earth coordinate and can provide stable patterns of body movements during exercises to enable our exercise recognition discussed in Section IV-C.

2) User-reference Alignment: We notice that using quaternion to align sensor reading from wearable coordinate to earth coordinate solves the different wearing orientation of wearable devices. Furthermore, we should also consider when people doing workout in gym with different facing directions.

Specifically, we convert the sensor readings from the mobile device coordinate to the human coordinate by using the quaternion-based rotation $p_h = q_{mh} p_m q_{mh}^{-1}$, where p_m

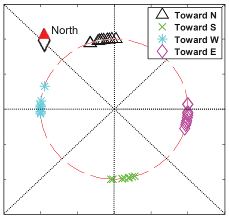


Fig. 8. Facing direction estimation of four running directions: toward North (N), South (S), West (W) and East (E).

and p_h is the sensor reading vector in the mobile device coordinate and the human coordinate respectively. q_{mh}^{-1} is the conjugate quaternion of q_{mh}, q_{mh} is the quaternion readings from the mobile device to the human coordinate, which can be calculated using Hamilton product: $q_{mh} = q_{he}^{-1} q_{me}$, where q_{me} is the quaternion reading from the mobile device coordinate to the earth coordinate, which can be obtained from the device directly. q_{he}^{-1} is the conjugate quaternion of q_{he}, and q_{he} is the quaternion readings from the human to the earth coordinate, which can be derived from the estimated facing direction.

More specifically, we can derive $q_{he} = [w, x, y, z]$ using the *Euler angles* in earth coordinate which is defined as:

$$\begin{cases} w = cos(\frac{\phi}{2})cos(\frac{\theta}{2})cos(\frac{\psi}{2}) - sin(\frac{\phi}{2})sin(\frac{\theta}{2})sin(\frac{\psi}{2}); \\ x = cos(\frac{\phi}{2})sin(\frac{\theta}{2})cos(\frac{\psi}{2}) + sin(\frac{\phi}{2})cos(\frac{\theta}{2})cos(\frac{\psi}{2}); \\ y = cos(\frac{\phi}{2})sin(\frac{\theta}{2})cos(\frac{\psi}{2}) - sin(\frac{\phi}{2})cos(\frac{\theta}{2})sin(\frac{\psi}{2}); \\ z = cos(\frac{\phi}{2})cos(\frac{\theta}{2})sin(\frac{\psi}{2}) + sin(\frac{\phi}{2})sin(\frac{\theta}{2})cos(\frac{\psi}{2}); \end{cases} \tag{5}$$

where rotation angles ϕ, θ and ψ are the row, $pitch$ and yaw respect to earth reference respectively as shown in Figure 7. We assume that people are running on the horizontal ground and therefore ϕ and θ are equal to zero and we only need to calculate facing direction ψ (i.e., yaw).

B. Facing Direction Estimation

We observe that in rest time and aerobic exercises, the direction of the user's arm swing is usually in line with the user's facing direction, suggesting that we can exploit the arm swing direction to estimate the user's facing direction. For anaerobic exercise, users can simply swing their arms for a few times to assist FitCoach for facing direction estimation.

In particular, FitCoach segments each arm swing using rep segmentation as described in Section IV-B, then converts the acceleration readings from mobile device's coordinate into earth coordinate as discussed in Section VI-A1. After conversion, we can double integrate the acceleration projected to the x and y axes in the earth coordinate to derive the moving distance of the arm along the x and y axes, respectively. In this work, we define the arm swing direction as the counter-clockwise rotation around the z-axis from y-axis in the earth coordinate (i.e., North direction), which is similar to the definition of yaw in *Euler angles*. We first calculate the included angle δ between the displacement of x-axis and y-axis caused by arm swing by using $\delta = |arctan\,(s_y/s_x)|$, where s_x, s_y are the distance accumulated from acceleration in x-axis and y-axis respectively by using *Trapezoidal rule* [12]. Note that δ is ranging from $0°$ to $90°$ and then we need to convert it from $0°$ to $360°$. Therefore, we need to decide the quadrant Q of arm swing direction, that is defined in Cartesian system where x and y are East and North in earth reference respectively, to convert it to ψ ranging from $0°$ to $360°$ as:

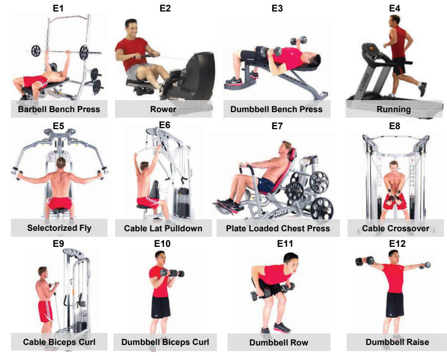


Fig. 9. Illustration of 12 types of exercises[1].

$$\psi = \begin{cases} 270° + \delta; & if\ Q = 1, \\ 90° - \delta; & if\ Q = 2, \\ 90° + \delta; & if\ Q = 3, \\ 270° - \delta; & if\ Q = 4, \end{cases} \tag{6}$$

where Q can be determined based on the order of maximum and minimum values (i.e., peak and trough) on x and y axes of accelerometer.

We evaluate the proposed facing direction estimation by asking a volunteer to run toward four different directions (i.e., north, south, east and west in earth reference). Figure 8 shows the 10-round estimation results for each direction. We find that the estimated results are along with the four running directions and good enough in FitCoach, the little bias is caused by the fact that people swing their arms naturally while running which is not perfectly stick to their facing directions.

C. Profiling Database Construction

When users start FitCoach for the first time, they are asked to build a profiling database for the exercise recognition by performing the particular types of exercises. FitCoach extracts the accl-based features as discussed in Section IV-C, and asks the user to manually label the corresponding exercise types. We note that FitCoach allows users to wear the wearable mobile devices with flexible facing orientation when constructing the profiling database, because the quaternion-based coordinate alignment always converts sensor readings to a coordinate system that has the fixed mapping relationship to the human coordinate during exercises.

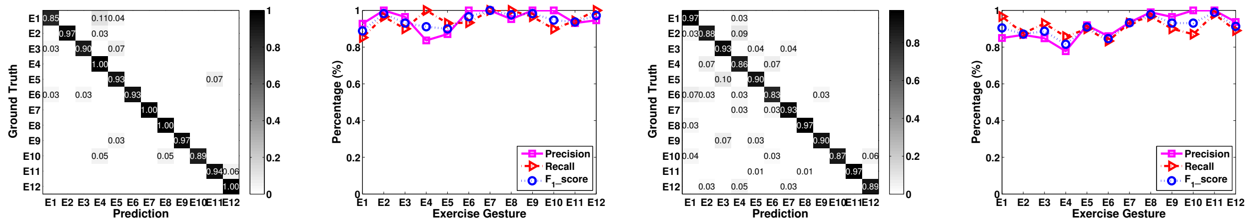
VII. Performance Evaluation

In this section, we first present the experimental methodology and metrics we used to evaluate FitCoach. We then evaluate the performance and robustness of FitCoach using both smartwatch and smartphone during people's fitness workout.

A. Experimental Methodology

1) Wearable Mobile Devices: We evaluate FitCoach with two types of wearable mobile devices (i.e., a smartphone of Samsung Galaxy Note 3 and a smartwatch of LG Watch Urbane). Both devices use Android and can collect sensor

[1]by courtesy of app *Fitness Buddy*.



(a) Confusion matrix, smartwatch (b) Precision/recall/F1 score, smartwatch (c) Confusion matrix, smartphone (d) Precision/recall/F1 score, smartphone

Fig. 10. Comparison of the performance of recognizing 12 exercises between using a smartwatch and a smartphone.

readings of accelerometer, gyroscope and quaternion vector. In our experiment, the participants are asked to wear the smartwatch on the wrist with their own wearing preferences and the phone is mounted on their upper arms using a jogging armband. During exercise, sensor readings are collected with the sampling rate of 100Hz. The ground truth of workout statistics are recorded by a volunteer.

2) Fitness Data Collection: We recruit 12 volunteers from colleagues, friends and students from research lab. Among them, 7 out of 12 go to gym regularly and the rest go to gym less frequently. For over a half year experiments, all 12 volunteers are asked to wear the smartwatch and smartphone simultaneously at the same arm, which is for the performance comparison between smartwatch and smartphone of the same exercise. In addition, a volunteer accompany with them to record the ground truth. Specifically, we study 12 different exercise types, as illustrated in Figure 9. The tested exercises include both anaerobic exercises, including weight machines and free weights, and aerobic exercises in which around 2 hours running is tested in both indoors (e.g., treadmill) and outdoors. In total, we collect over 5000 repetitions of 12 types of exercises involving 12 participants.

B. Evaluation Metrics

We use the following metrics to evaluate FitCoach:

Precision. Given N_e reps of a exercise/ gesture type e in our collected data, precision of recognizing the exercise type e is defined as $Precision_e = N_e^T/(N_e^T + M_e^F)$, where N_e^T is the number of instances collectedly recognized as exercise e. M_e^F is the number of sets corresponding to other exercises that are mistakenly recognized as exercise e.

Recall. Recall of the exercise type e is defined as the ratio of the reps that are correctly recognized as the exercise e over all reps of exercise type e. which is defined as $Recall_e = N_e^T/N_e$.

F1-score. F1-score is the harmonic mean of precision and recall, which reaches its best value at 1 and worst at 0. In our multi-class scenario, the F1-score for a specific gesture e was defined as $F_1^{(e)} = 2 \times \frac{precision_e \times recall_e}{precision_e + recall_e}$.

Rep Detection Rate. Given all reps of an exercise type e, rep detection rate is defined as the ratio of the number of detected reps of e over all reps of e the user performed.

C. Workout Recognition Using Smartwatch

We first evaluate the performance of FitCoach on exercise recognition using smartwatch. Figure 10(a) shows the confusion matrix of the recognizing exercise types by using smartwatch in FitCoach. An entry M_{ij} denotes the percentage between the number of exercise i was predicted as gesture j and the total number of i. The average accuracy is 95% with standard deviation 5% over all 12 types of exercises. We find that recognizing results from $E1$ and $E10$ are relatively low, which are 85% and 89% respectively. This may be caused by some volunteers who go to gym less frequently and cannot maintain the exercise in a correct form for all reps. For example, $E10$ (i.e., Dumbbell Biceps Curl) is free weight exercise and some volunteers may not maintain their arm within a fixed space all the time. For exercise $E1$ (i.e., Barbell Bench Press), some volunteers easily perform too fast or too slow depending on the weights.

In addition, Figure 10(b) presents the precision, recall and F_1 score for each exercise type, respectively. The average value of precision, recall and F_1 score of each exercise are all around 95%. Although the recall of exercise $E4$ (i.e., running) is 100%, we observe that it has the lowest precision among all 12 exercises, which indicates other exercises are more likely to be mistakenly classified as this exercise. This may be caused by the fact that arm swings are naturally moving in space and some volunteers freely perform some type of exercise too fast which also involve all axes sensor readings. The above results support that FitCoach can extract accurate information for exercise type recognition through wrist-worn smartwatch.

D. Workout Recognition Using Smartphone

We then evaluate workout recognition by using smartphone since arm-mounted phone have been widely used in people's daily exercise. We present the results from smartphone in Figure 10 (c) and Figure 10 (d). Results show 91% average recognition accuracy for exercise recognition. We find exercise $E4$ still has the lowest precision which is consistent with the results collected from smartwatch since the volunteers wear smartwatch and smartphone on the same arm to make fair comparison.

Comparison between Smartwatch and Smartphone. FitCoach presents high accuracy of workout recognition for both smartphone and smartwatch. Comparing results between smartwatch and smartphone, we found that results obtained from smartwatch are better than results from smartphone. The average recognition accuracy of smartwatch is 95% whereas smartphone has a 91% average recognition accuracy. This observation is due to the fact that for exercise recognition, the space scope of the arm gesture trajectories was constrained by the machine for some exercise and most of the exercises

Fig. 11. Detection rate of exercise repetitions by using smartwatch.

require users to use their hands to grab and therefore the smartwatch on the wrist are close to hand and reflect more similar movement as machine or dumbbell.

E. Rep Detection Rate

Finally, we evaluate FitCoach by showing our detection rate for exercises. For workout exercise detection, the average detection accuracy reaches 99%. The lowest detection rate occurs at running exercise $E4$ (i.e., step detection) on a treadmill but it still achieves around 95% detection accuracy as shown in Figure 11. Such relative low detection rate of running exercise is cased by occasionally holding on the handrails or wiping perspiration while running. The above results show that FitCoach can accurately detect reps, and such high detection rate supports that fine-grained statistical information provided by FitCoach is reliable.

VIII. Conclusion

In this work, we propose FitCoach, an integrated mobile solution that can conduct systematic fitness monitoring and provide performance review based on a single off-the-shelf wearable device (e.g., wrist-worn wearables or arm-mounted smartphones). FitCoach has the capability to perform fine-grained exercise recognition including exercise types, the number of sets and repetitions by using inertial sensors from wearable devices without user involvement. Two novel metrics, exercise form score and workout review plane, are developed to provide effective review and recommendation for achieving effective workout and preventing injuries. To ensure the system accuracy and robustness, FitCoach uses the earth/human coordinate system to align and integrate sensor readings from various device orientations. Extensive experiments involving 12 participants doing workout for over half a year time period demonstrate that FitCoach successfully takes one step forward to provide the integrated fitness monitoring system with over 90% workout analysis accuracy. By integrating other existing sensors such as shoe sensors [4] and ankle-based belt, FitCoach can be extended to monitor non-arm based exercises. In addition, FitCoach can further reduce the energy consumption by utilizing location information. The system only needs to start sampling when detecting gym or fitness center nearby through the assistant of GPS and we left this part in our future work.

IX. Acknowledgments

This work is supported in part by the NSF grants CNS1514436 and CNS1409767.

References

[1] Fitbit. https://www.fitbit.com/.
[2] Garmin watch. http://www.garmin.com/en-US/.
[3] Gymwatch. https://www.gymwatch.com/.
[4] Nike+ sensor. https://secure-nikeplus.nike.com/plus/products/.
[5] Sworkit. http://sworkit.com/.
[6] J. Arney. You should be in pictures! experts share tips on how to produce a quality fitness video for any purpose. *IDEA Fitness Journal*, 2(1):86–90, 2005.
[7] K.-h. Chang, M. Y. Chen, and J. Canny. Tracking free-weight exercises. In *Proceedings of the 9th international conference on Ubiquitous computing (ACM UbiComp)*, 2007.
[8] H.-T. Cheng, F.-T. Sun, M. Griss, P. Davis, J. Li, and D. You. Nu-activ: Recognizing unseen new activities using semantic attribute-based learning. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services (ACM Mobisys)*, 2013.
[9] J. R. Deller, Jr., J. G. Proakis, and J. H. Hansen. *Discrete Time Processing of Speech Signals*. Prentice Hall PTR, 1st edition, 1993.
[10] H. Ding, L. Shangguan, Z. Yang, J. Han, Z. Zhou, P. Yang, W. Xi, and J. Zhao. Femo: A platform for free-weight exercise monitoring with rfids. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (ACM Sensys)*, pages 141–154, 2015.
[11] T. Hao, G. Xing, and G. Zhou. Runbuddy: a smartphone system for running rhythm monitoring. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*, 2015.
[12] B. Lindberg. On smoothing and extrapolation for the trapezoidal rule. *BIT Numerical Mathematics*, 11(1):29–52, 1971.
[13] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng. Tracking vital signs during sleep leveraging off-the-shelf wifi. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM Mobihoc)*, pages 267–276, 2015.
[14] M. D. Mifflin, S. T. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. Koh. A new predictive equation for resting energy expenditure in healthy individuals. *The American journal of clinical nutrition*, 51(2):241–247, 1990.
[15] F. Mokaya, R. Lucas, H. Y. Noh, and P. Zhang. Myovibe: Vibration based wearable muscle activation detection in high mobility exercises. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*, pages 27–38, 2015.
[16] A. C. of Sports Medicine et al. *ACSM's guidelines for exercise testing and prescription*. Lippincott Williams & Wilkins, 2013.
[17] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis. RisQ: recognizing smoking gestures with inertial sensors on a wristband. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services (ACM MobiSys)*, pages 149–161, 2014.
[18] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang. Smartphone based user verification leveraging gait recognition for mobile healthcare systems. In *Proceedings of the 10th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (IEEE SECON)*, 2013.
[19] Z. Ren, J. Yuan, J. Meng, and Z. Zhang. Robust part-based hand gesture recognition using kinect sensor. *IEEE Transactions on Multimedia*, 15(5):1110–1120, 2013.
[20] E. Thomaz, I. Essa, and G. D. Abowd. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*, pages 1029–1040, 2015.
[21] V. N. Vapnik and V. Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998.
[22] D. Vlasic, R. Adelsberger, G. Vannucci, J. Barnwell, M. Gross, W. Matusik, and J. Popović. Practical motion capture in everyday surroundings. In *ACM Transactions on Graphics (TOG)*, volume 26, page 35, 2007.
[23] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking (ACM Mobicom)*, pages 617–628, 2014.