# WHEN VIRTUAL REALITY MEETS IOT IN THE GYM: ENABLING IMMERSIVE AND INTERACTIVE MACHINE EXERCISE

By

Md Fazlay Rabbi

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering - Master of Science

2016

**ABSTRACT**

WHEN VIRTUAL REALITY MEETS IOT IN THE GYM: ENABLING IMMERSIVE AND INTERACTIVE MACHINE EXERCISE

By

Md Fazlay Rabbi

The advent of head mounted displays (HMDs) such as Oculus Rift, Samsung Gear VR, and Microsoft HoloLens are turning immersive virtual reality (VR) into reality. As one of its most compelling applications, we envision that VR will revolutionize the personal fitness experience in our daily lives. Towards this vision, we present *JARVIS*, a novel immersive exercise tracking applications enabled by miniature IoT sensing devices combined with a mobile HMD device. By attaching IoT sensing devices on any gym machines, JARVIS continuously tracks exercise progress and assesses exercise quality in real-time, enabling effective interaction with machines. By converting the captured exercise progress and quality information to VR inputs, it creates an immersive exercise experience with a virtual avatar to guide machine exercises. We have conducted extensive experiments to validate the performance of JARVIS. It achieves 97.96% of repetition segmentation accuracy without knowing the current exercise type, as well as 99.08% of exercise type recognition accuracy in 1.22 repetitions on average. We have also conducted a real-world deployment study to examine the efficacy of the proposed platform. By analyzing muscle activities using high-fidelity clinical surface Electromyography (sEMG) sensors, our results indicate that our VR/IoT-based platform could provide an engaging and effective guidance to exercisers for their strength training.

To my mom (Fatema Khatun) and dad (Ansar Ali Mia) for their unconditional love and support throughout my education.

# ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere appreciation to my advisor, Dr. Mi Zhang, for his continuous support, help, patience and encouragement throughout my master study. I would also like to thank Dr. Taiwoo Park from the Department of Media and Information for his help, support and motivation throughout my thesis work. This thesis would not have been possible without the guidance of Dr. Mi Zhang and Dr. Taiwoo Park. I am deeply grateful to Dr. Youngki Lee from Singapore Management University for his motivation and insightful comments about this work.

My sincere gratitude also goes to my committee member, Dr. Daniel Morris. Dr. Morris was very generous with his time, ideas and assistance.

Special thanks go to my colleagues in the Ubiquitous and Mobile Sensing Laboratory who have supported me in my research work through reviews, discussions and meetings. I am thankful to my friends (in no particular order) Sayeed Hyder, Shams Reaz, Saif Muhammad Imran, Sheikh Mohammad Shavik, Proteek Chandan Roy, Afibuzzaman Anik and Kamran Ali for their support and making my life wonderful here in East Lansing.

Last but not least, I am deeply grateful to my family for their tremendous support and encouragement.

# TABLE OF CONTENTS

# LIST OF TABLES

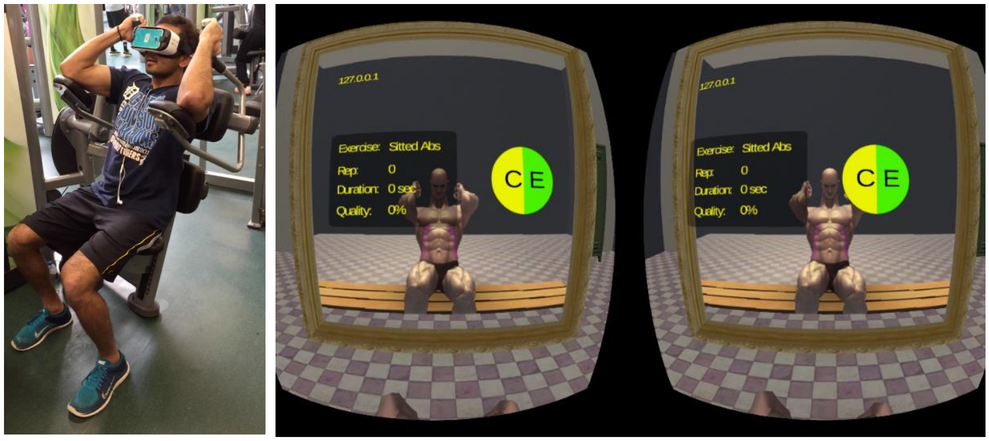
# LIST OF FIGURES

# Chapter 1

# Introduction

The promise of immersive virtual reality (VR) is starting to look very real with the emergence of head mounted displays (HMDs) such as Oculus Rift [5], Samsung Gear VR [6], and Microsoft HoloLens [3]. Immersive VR soaks a user in a computer-generated simulated environment that changes naturally through reflecting the movements of the user. An engaged immersive virtual experience is thus realized by employing sensing technologies that capture the user's movements and use those information to update the sensory stimuli presented to the user via a HMD to create an illusion of being immersed in a virtual environment in which they can interact [27]. Given its unique capability of enabling such engaged immersive virtual experience, immersive VR has been regarded as a technology that has the significant potential to revolutionize a wide range of industries such as entertainment, education, fitness and healthcare.

In this thesis, we propose an innovative *immersive and interactive exercise assistant*, named *JARVIS*, enabled by synergistic adoption of two emerging technologies, i.e, Head-mounted VR devices and minature IoT sensors. Today, workout in the gym has become an important part of people's modern lifestyle [28]. However, working out on the stationary exercise machines in the gym quickly make exercisers feel easily bored [15, 18]. Moreover, novice exercisers are hardly aware if they are using the right set of muscles, if their speeds of exercises are adequate, etc. without help of professional trainers. This prevents exercis-




(a) (b)

Figure 1.1: Example usage: (a) a user trying JARVIS, (b) stereoscopic immersive VR screen of virtual avatar.

ers from making steady progress, and eventually makes exercisers lose their interests and motivation.

*JARVIS* is the first-of-its-kind exercise assistant system to enable immersive and interactive gym exercise experience. In a nutshell, JARVIS senses exercise types and movements of a exerciser using a miniature portable sensor device, assesses exercise quality and progress in real-time, and provide the rich, immersive feedback through a VR headset. For example, it shows a user a virtual avatar scene with her virtual body, indicating proper way of exercise, along with rich real-time exercise progress and quality information. In this way, JARVIS aims at serving a virtual avatar to guide machine exercises in a highly interactive manner and also creates a truly immersive gym exercise experiences.

A number of unique technically and usability challenges arises to fully realize the vision of JARVIS. First of all, it needs to sense the type of exercises and movements of exercisers in a real time. Since most of the gyms are not yet instrumented with sensing capability, we aimed at achieving this by a single tiny portable sensor that the exerciser carries. Second, JARVIS need to meet very tight performance requirements in terms of latency and accuracy.

Low latency and high accuracy are very important to enable usable and highly immersive experiences. Third, we need to create smooth VR scenes without making the head-mounted VR device heated too much. Exercisers easily feel hot while lifting the machines, and even moderate heat caused by the VR device could easily hinder pleasant exercise experience. Finally, we need to design a usable, immersive, and effective VR avatar user interface.

JARVIS addresses these challenges in several ways. First, it employs the concept of machine-wearables, that is, temporarily attaching a tiny sensor IoT tag on exercise machines, thereby obtaining clean sensor signal without pre-installed infrastructure. Second, we find a combination of tag locations showing the best exercise type recognition accuracy, as well as utilizing repetitive nature of exercise to further increase the accuracy. Third, we propose a smart guide for UI developers to adjust trade-off between visual quality and computation load. Lastly, we design an immersive and effective virtually-reconstructed user interface, by fully utilizing the potential of VR HMD device.

We have shown the effectiveness of JARVIS through extensive experiments on its technical components as well as field studies on its usability. Our results show that JARVIS could quickly and accurately segment repetitions (97.96%) and detect exercise type (99.08% in 1.22 repetitions on average), with a tiny IoT device. Its smart assist for UI developers also helps in preventing the VR HMD device from overheating. Our field case study conducted over 10 exercisers shows the efficacy of the virtual avatar enabled by the system. JARVIS showed significant differences in enjoyment, perceived competence, and usefulness, compared to a typical exercise setting, and the participants generally liked the VR user interface components. In addition, we quantitatively analyze muscle activities of exercisers using high-fidelity clinical surface Electromyography (sEMG) sensors, and our results indicate that virtual muscle highlighting provides a significant difference ($p < .01, p < .05$) in muscle activation, compared to a typical exercise training, implying the potential of effective guidance for strength training.

# Chapter 2

# Background

## 2.1 Motivation and Design Objectives

### 2.1.1 Motivating Scenario

David has been thinking of building his muscles using stationary workout machines in the gym for recent couple years. However, he failed to adhere to regular strength training schedules. He once worked with personal trainers, but they could not be with him all the time. Most of the existing mobile fitness logging apps required him to manually log exercise records. Although there are a few automatic fitness monitoring apps, they could only summarize the exercises after the exercise session, which was not very helpful for him *during* the training.

David begins to try JARVIS exercise tracking platform. He brings a mobile VR HMD and a small sensor tag to any gym. He wears the HMD and puts the sensor tag on a biceps curl machine. After a few number of trial repetitions, JARVIS detects the type of the machine and quickly displays a virtual trainer scene, designed for the biceps curl. JARVIS also provides immersive and interactive VR environment showing a "virtual David" following his exercise movements, and provides real-time exercise progress and quality information. Moreover, JARVIS highlights the required muscle groups corresponding to the current exercise being performed and his personal goal of strength training, which helps him easily give more focus on the muscle groups. Once David finishes biceps curl, JARVIS turns off the virtual exercise scene and switches into outside pass-through camera, to let David navigate to the next machine. With JARVIS, David can go to the gym with automatically-generated training plans, and an automated real-time immersive exercise tracking application which makes him more confident in his strength training.

### 2.1.2 Design Objectives

Based on the motivating scenario, the design of JARVIS aims to achieve the following goals:

- *Fully Interactive and Immersive Exercise Experience:* JARVIS is designed to create virtual reality machine exercise training scene with a *virtual avatar* to provide users an immersive and interactive machine exercise experience.

  One of the key roles of a virtual avatar is to interact with a user by providing guidance on the performed exercise in real time. To achieve this goal, JARVIS first provides a scene in VR HMD, showing a virtual body of the user, following the user's movement, to intuitively guide her exercises. It also aims to provide user's exercise progress and quality information in real time.

- *Comprehensive Machine Exercise Analytics:* JARVIS is designed to provide *comprehensive analytics* of the machine exercises to users. Existing work in exercise sensing systems focus on detecting exercise sessions, counting exercise repetitions as well as recognizing exercise types [21, 9]. Although repetitions and types are useful for tracking exercise progress, providing feedback to users about the quality of their exercises is more important to novice and intermediate users. As such, JARVIS aims to provide a comprehensive machine exercise analytics by tracking the exercise progress and quantifying the quality of the performed machine exercises.

- *Wearable for Machines:* JARVIS is designed to be "wearable" for exercise machines. Existing work use smartphones and wearable devices such as smartwatches and armbands (i.e., human-wearables) to track free weight and body weight exercises [21, 22]. However, in the domain of machine exercises, using sensing devices temporarily attachable to machines (i.e., *machine-wearables*) has two significant advantages over using human-wearable devices. First, *machine-wearables* can capture abdominal and lower limb machine exercises that human-wearables fail to capture. Second, *machine-wearables* only capture machine's constrained movements, thereby providing much cleaner motion data than human-wearables. In contrast, human-wearables capture body movement noises as well as non-exercise body movements between exercise sessions, which requires significant signal processing efforts to filter out. JARVIS aims to leverage these advantages to precisely and accurately monitor and detect exercise context information.

- *Universal Sensing Platform:* JARVIS is designed to provide a *universal sensing platform* that can be eventually used for any exercise machine in a *plug and play* manner. Pioneer work in powering exercise machines with sensing capabilities explored customized instrumentation of exercise machines by integrating different types of sensing devices into different exercise machines [25]. This approach requires considerable efforts from machine manufacturers in modifying exercise machines, leading to significant increase of costs of the exercise machines. In contrast, JARVIS aims to reduce the burden of machine manufacturers by developing a uniform sensing device that can provide sensing capability to any exercise machine without customized modification. We envision that in the future, every exercise machine will have a standardized slot/interface for the uniform sensing device to plug in. The exerciser will plug out the device after the exercise. As such, JARVIS acts as a personal platform that tracks individual's exercises on any machine.

To the best of our knowledge, no existing exercise sensing systems meet all of the above design goals. This motivates us to design JARVIS to fill this critical gap.

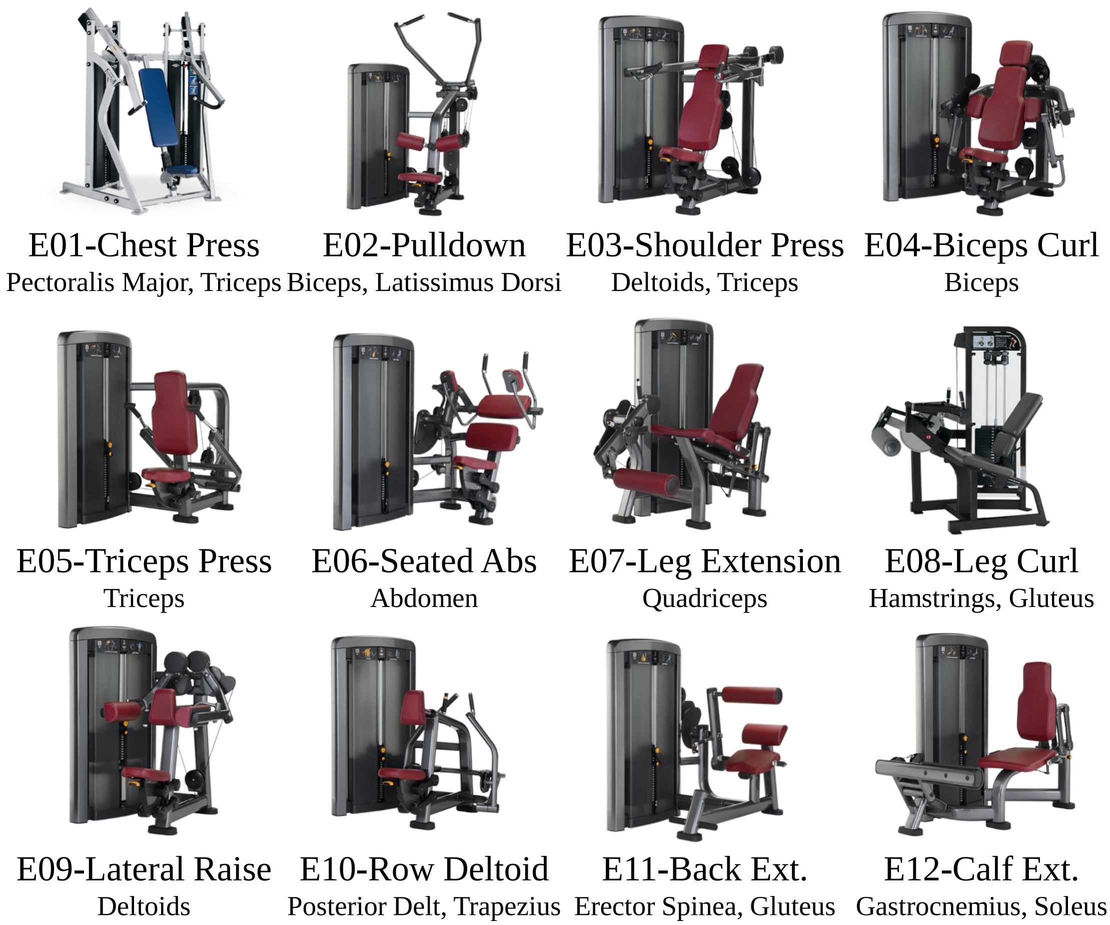


Figure 2.1: The 12 target machine exercises.

### 2.1.3 Targeted Machine Exercises

In this work, we target twelve machine exercises recommended by the resistance training guide for healthy adults from the American College of Sports Medicine (ACSM) [8]. These exercises represent the most common machine exercises that target different muscle groups on the body. Each exercise uses a dedicated machine to train the specific muscle group. Figure 2.1 illustrates the 12 targeted machine exercises and corresponding muscle groups.

# Chapter 3

# Related Work

Immersive VR has undergone a transition in the past few years that has taken it out of the realm of expensive toy and into that of functional technology [10]. As one of its most important applications, we envision that immersive VR will fundamentally change people's exercise experiences.

In the past decade, the academia and industry have achieved significant success in developing wearable and mobile sensing systems for tracking aerobic exercises such as walking, jogging and running. Recently, research focus has shifted to sensing anaerobic exercises (i.e., muscle strength training exercises). In general, the muscle strength training exercises can be grouped into three categories: free weight exercises, body weight exercises, and machine exercises. Most of the existing work focused on sensing free weight and body weight exercises. For example, in [9], Chang *et al.* demonstrated the feasibility of using two accelerometers by wearing one on the hand and the other on the waist to track free weight exercises. In [21], Morris *et al.* developed RecoFit that used a single inertial sensor worn on the right forearm of an individual to monitor both free weight as well as body weight exercises. In [23], Muehlbauer *et al.* leveraged the accelerometer and gyroscope inside a smartphone and wore the smartphone on the arm to track upper body exercises. The fundamental difference between these works and JARVIS is that they use smartphones or custom-made sensing devices worn on the human body to monitor free weight and body weight exercises. In comparison, our work uses miniature IoT-based sensing devices "worn" on the stationary gym machines to track machine exercises. Finally, in [11], Ding *et al.* developed FEMO, a RFID-based sensing system for free weight exercise monitoring. Our work is similar to FEMO in the sense that we both attach sensors onto exercise instruments (FEMO attaches RFIDs to dumbbells). However, since FEMO uses RF signals to track dumbbell movements, its performance suffers from the interferences caused by the movements of other exercisers in the gym. In contrast, JARVIS uses accelerometer and gyroscope to capture machine movements caused by the exerciser' exercise-related body motions, which are not interfered by the movements of other exercisers. More importantly, JARVIS leverages immersive VR to create controllable 3D stimulus environments as well as an engaged virtual avatar to guide exercisers in a highly interactive way that was not previously possible using existing approaches.

# Chapter 4

# JARVIS System Overview

Figure 4.1 illustrates the system architecture of JARVIS. As illustrated, JARVIS consists of two devices: a miniature IoT sensing device that is attachable to gym machines to track machine exercises; and a VR HMD that processes the sensor data and visualizes the computer-generated simulated environment as well as the exercise information to an exerciser via the HMD UI in real time.

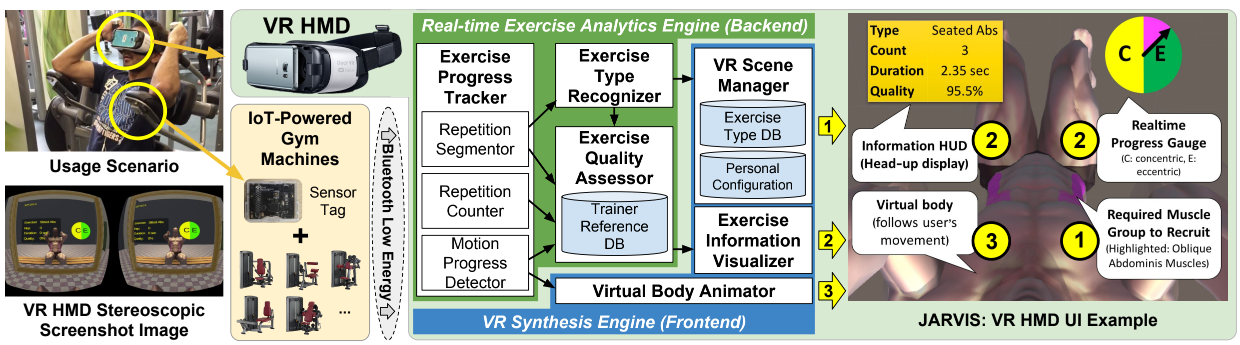



Figure 4.1: The system architecture of JARVIS.

As the core of JARVIS, the *Real-time Exercise Analytics Engine* and the *VR Synthesis Engine* run inside the VR HMD at the backend and the frontend respectively.

At the backend, *Real-time Exercise Analytics Engine* retrieves the sensor data from the IoT sensing device and analyzes the sensor data. Specifically, *Real-time Exercise Analytics Engine* consists of three major components: *Exercise Progress Tracker*, *Exercise Type Recognizer*, and *Exercise Quality Assessor*. The role of *Exercise Progress Tracker* is to segment the streaming sensor data into individual exercise repetitions (*Repetition Segmentor*), count the repetition numbers (*Repetition Counter*), and track the progress of exercise within each repetition (*Motion Progress Detector*). Given the segmented repetitions, *Exercise Type Recognizer* detects the exercise type of each repetition while *Exercise Quality Assessor* provides a quantitative evaluation on the performed exercise repetition by comparing it with the guide models from professional trainers.

At the frontend, *VR Synthesis Engine* synthesizes the immersive computer-generated simulated gym exercise environment with a virtual body of the exerciser. It also provides real-time exercise analytics based on the exercise information from the backend *Real-time Exercise Analytics Engine*. Specifically, *VR Synthesis Engine* consists of three major components: *VR Scene Manager*, *Exercise Information Visualizer*, and *Virtual Body Animator*. The role of *VR Scene Manager* is to initiate a virtual exercise scene corresponding to the current exercise type, recognized by *Exercise Type Recognizer* in the backend, while highlighting target muscle group based on exercise and personal profile. In addition, it provides visual quality and computation load trade-off summary to UI developers. *Exercise Information Visualizer* generates a head-up display in a virtual space and delivers exercise progress and quality information to a user in real time. *Virtual Body Animator* makes a virtual body to move following the user's movement, based on the real-time progress information provided by *Motion Progress Detector*.

In the following two sections, we describe the *VR Synthesis Engine* and the *Real-time Exercise Analytics Engine* in details.

## 4.1 VR Synthesis Engine

### 4.1.1 Automatic Scene Management

Once a user begins exercising, *VR Scene Manager* automatically initiates a virtual exercise scene, based on the exercise type recognized by the *Exercise Type Recognizer*. Figure 4.1 (on the right side) shows an example scene for a seated abs machine. For user convenience, while a user is not exercising, the manager shows outside (i.e., pass-through) using the camera placed at the back of the VR HMD.

#### 4.1.1.1 Muscle Highlighting

In the scene, *VR Scene Manager* visually highlights target muscle groups, corresponding to exercise type (e.g., seated abs) and personal goal (e.g., muscle shaping, better spine structure), retrieved from exercise/personal goal preset databases. For example, in the UI example of Figure 4.1, the app paints oblique abdomens muscles purple, at which users should give their attention during exercise. This feature hypothesizes visual highlight of target muscle will lead to greater Mind-muscle connection (MMC) [30]. MMC is a practical term denoting a strategy to give attentional focus to consciously direct neural drive to the muscle, usually achieved through imagination [30]. Increased MMC leads to greater muscle activation, which potentially increases muscle protein accretion [32, 33].

#### 4.1.1.2 Smart Visual Quality Guide

One important feature of *VR Scene Manager* is to provide UI developers useful suggestions to achieve higher perceived VR rendering quality, while reducing computation workload of


Figure 4.2: Example of three different quality options.

CPU and GPU. Along with the profiler tool of Unity[1], the system provide UI developers computation load profile and suggestions to give quality graphics (i.e., detailed models with more polygons and shader effects) on the target muscle group, while drawing the remaining body parts using moderate to low number of polygons and/or less expensive shader effects. This development-side suggestion aims to enable higher perceived visual quality, while reducing energy consumption and heat generated by VR HMD. For example, Figure 4.2 shows a balanced option of quality and computation load (in the middle), by mixing a standard shader and high quality shader (e.g., a bumped specular shader with lightmap). Later we evaluate how this approach contributes to reduced heat generation of the VR HMD, which is beneficial for exercisers' experience as well as preventing overheat termination of the VR application.

### 4.1.2 Real-time Virtual Body Animation

Virtual Body Animator generates a virtual body of a user, that follows the user's movement during the exercise, based on the real-time progress information provided by *Motion Progress Detector*. It enables the user to intuitively understand the pace and progress of the current

[1] A widely-used game and interactive app development framework.

repetition. After several rounds of pilot test in development phase, we put a mirror in front of the virtual body, to better show users their virtual body and its movement, which is turned out to be effective in our case study.

### 4.1.3 Informative Head-Up Display (HUD)

*Exercise Information Visualizer* collects real-time exercise information from the backend architecture and displays the information on the VR screen. It employs a HUD interface in a fixed location, indicating rich exercise information including repetition count, pace, phase information, and quality. It also shows whether the pace of exercise is too fast or slow, in comparison with the recommendation of American College of Sports Medicine (ACSM) [8]. The recommendation provides pace guidelines, for example, for fast (less than 2 seconds per repetition) and moderate (2-4 seconds) pace of strength training. Also, it provides pace breakdown for two phases in a repetition: eccentric and concentric phases, for exercisers to maximize their performance [17].

## 4.2 Real-time Exercise Analytics Engine

### 4.2.1 Data Acquisition and Preprocessing

We use CC2650STK sensor tag device developed by Texas Instruments (TI) [7] to equip exercise machines with sensing capabilities. SensorTag is TI's state-of-the-art IoT device that integrates high-performance sensors and low-power wireless communication in a miniature form factor. In this work, we use the on-board 3-axis accelerometer and 3-axis gyroscope with a sampling rate of 10Hz to capture machine exercises. After we apply a moving average filter with length of 10 to suppress high frequency noises.

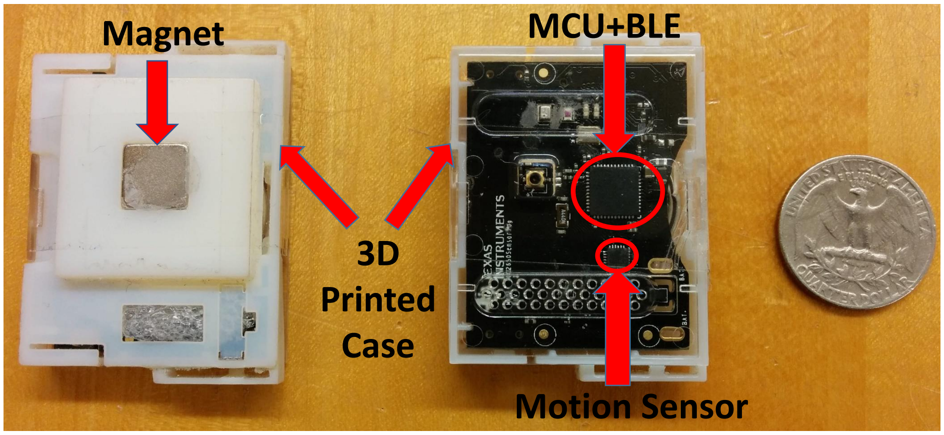



Figure 4.3: The miniature SensorTag device and a customized 3D printed case with an embedded magnet.

To help exercisers easily attach the sensor tag to exercise machines, we have designed and 3D printed a plastic case ($1.79 \times 2.64 \times 0.55$ inch) with an embedded magnet to host the sensor tag. With the magnet, the tag can be easily yet firmly attached to exercise machines (See Figure 4.3).

### 4.2.2 Exercise Progress Tracking

#### 4.2.2.1 Repetition Segmentation and Counting

The goal of repetition segmentation is to segment the streaming sensor data so that each segment contains one complete repetition of the performed machine exercise. Since a user can place the sensor tag on exercise machines in different ways which leads to different orientations, one straightforward scheme is to derive the orientation-independent acceleration magnitude signal and then apply peak detection on top of it to segment exercise repetitions. However, such scheme is unsuitable in the context of machine exercises because different machine exercises will have different numbers of peaks and valleys in each repetition. As an example, Figure 4.4(a) and Figure 4.4(c) illustrate the three axes of the accelerometer signal as well as the corresponding acceleration magnitude signal of three repetitions of *Pulldown* and *Seated Abs* respectively. As shown, within each repetition segment, the acceleration magnitude signal of *Seated Abs* has one peak and two valleys while *Pulldown* has three peak and two valleys. Without knowing the machine exercise type a prior, by simply applying peak detection, one repetition can be split into multiple segments.

In this work, we design a Principal Component Analysis (PCA) based scheme to segment repetitions. The key observation behind our scheme is also illustrated in Figure 4.4. Specifically, Figure 4.4(b) and Figure 4.4(d) illustrate the first principal component (PC) extracted from the three axes of the accelerometer signal of *Seated Abs* and *Pulldown* respectively. We observe that even though the exercise type is different, each repetition intersects the mean crossing line of the first PC signal exactly twice. The same observation holds true for all the 12 targeted machine exercises. This is because one repetition of any type of the targeted machine exercises consists of one concentric phase (i.e., muscle shortening) and one eccentric phase (i.e., muscle lengthening). And the first PC reliably captures both two phases.

Based on the key observation, our PCA based repetition segment scheme first extracts the first PC of the 3-axis accelerometer data and finds the mean crossing point of the first PC. Second, our scheme finds out whether the first PC is going downward or upward at the first mean crossing point. If going downward, the lowest minima between every non-overlapping pair of mean crossing points is the peak of the repetition and the two closest maxima on the left and right side of the two mean crossing points are the start and end point of the repetition. If going upward, the highest maxima between every non-overlapping pair of mean crossing points is the peak of the repetition and the two closest minima on its left and right are the start and end of the repetition.

After a new repetition is segmented, the number of segmented repetitions will be counted

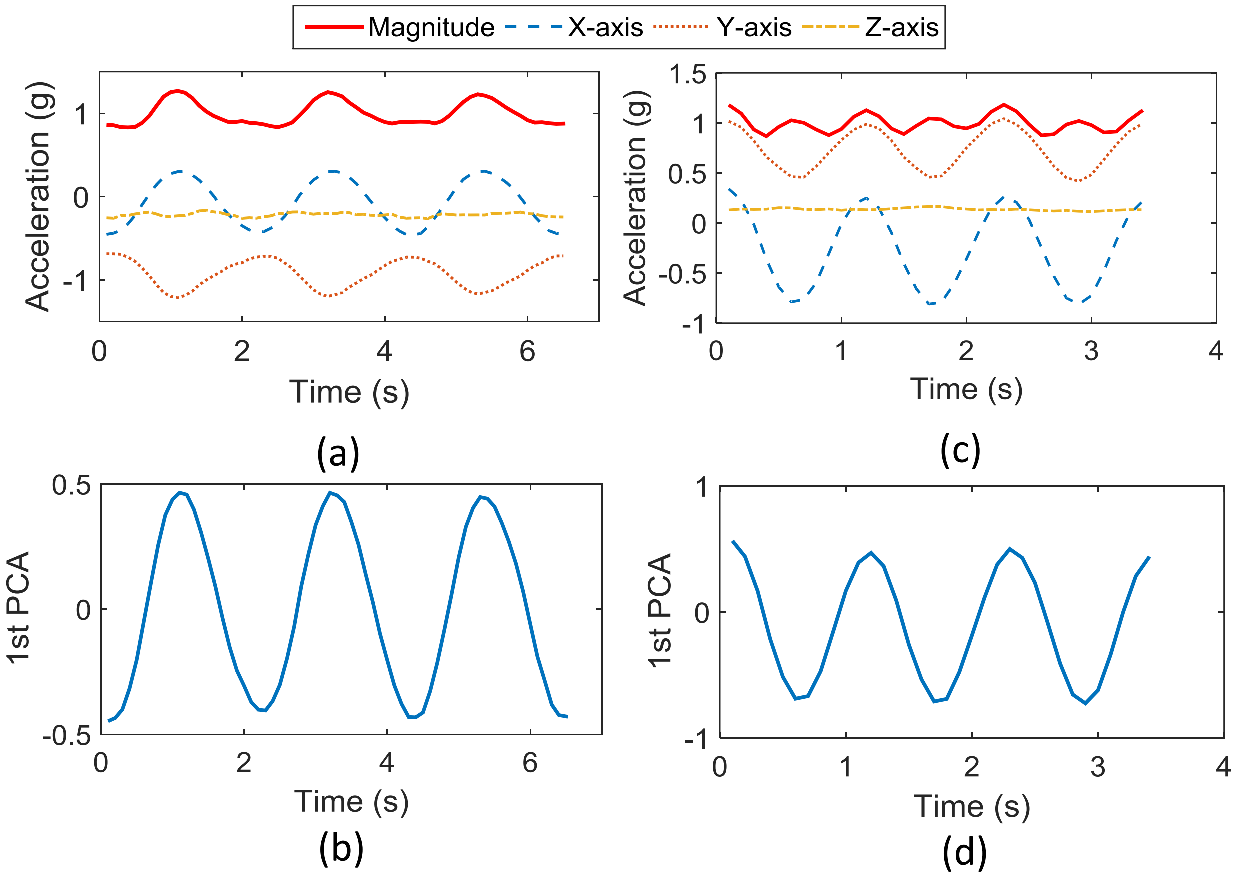



Figure 4.4: The illustration of the principle of the repetition segmentation algorithm. (a) and (c): 3-axis accelerometer data and the corresponding acceleration magnitude signal of three repetitions of Pulldown (a) and Seated Abs (c). (b) and (d): the first principal component (PC) extracted from the three axes of the accelerometer signal of Pulldown (b) and Seated Abs (d).

and updated in real time.

#### 4.2.2.2 Motion Progress Detection

To enable engaged interaction between the exerciser and the virtual avatar, JARVIS needs to track the progress of the exerciser within each repetition in real time. However, providing the exact progress status within each repetition in real time is not possible because the exact progress status can only be available after seeing the complete repetition. As an alternative, we design a motion progress estimation algorithm to provide a reasonable estimation about the progress status within each repetition in real time. Specifically, our algorithm uses the values of the first PC signal to estimate the progress status. The progress status starts from 0% and ends at 100% with a step of 10%. The specific values of the first PC signal that correspond to those status percentages are obtained by previously seen repetitions as training data. During real-time progress tracking, if the value of the first PC falls between two status percentages, the higher percentage will be reported as the estimated progress status.

### 4.2.3 Exercise Type Recognition

After segmenting the machine exercises into repetitions, the second stage of the machine exercise analytics engine is to identify the type of the machine exercise within each repetition. Due to the different mechanical constraints of exercise machines, each type of machine exercise has a certain form. Therefore, we leverage this observation and frame the machine exercise identification problem as a classification problem. As explained before, exercisers could place the sensor tag on exercise machines in different ways which leads to different orientations. To make our classification algorithm orientation-independent, we compute the magnitude of the three dimensional accelerometer data as well as the magnitude of the three dimensional gyroscope data within each repetition. Based on these two magnitudes, we have extracted a total number of 28 features which have been proven to be effective for activity recognition [34]. Table 4.1 summarizes the list of features. Finally, we stack the extracted features into a feature vector and import the feature vector into the linear kernel Support Vector Machine for classification.

#### 4.2.3.1 Sensor Location

To increase recognition accuracy, we leverage the fact that the sensor tag can be attached on any steel part of exercise machines. We collect motion data from two locations from each machine, and find the best combination of locations across the machines, by trying all possible location combinations. In this work, we employ 12 machines, and thus the number of all possible sensor location combinations is 4096 ($2^{12}$). Our criteria in choosing two locations are as following: (1) locations should be easily accessible by users, and (2) two locations on one machine should show different range of motion. We tagged locations with larger range of motion as 'L' and ones with smaller range as 'S', and later compare accuracies using different location combinations.

#### 4.2.3.2 Feature Selection

To find the best feature set providing the highest accuracy, we utilize the Sequential Floating Forward Selection (SFFS) feature selection algorithm [13] to identify a minimal subset of features achieving the best classification accuracy.

#### 4.2.3.3 Session-wise Voting

To maximize the recognition accuracy, we utilize a 'voting' scheme across repetitions in the same session. That is, we take the majority of recognition results from all repetitions in one session. Moreover, considering the requirement of on-line, real-time recognition, we later evaluate how many repetitions are required to achieve the highest accuracy.

| | | |
|---|---|---|
| Mean | Median | Standard Deviation |
| Variance | Skewness | Kurtosis |
| Energy | Interquartile Range | Spectral Entropy |
| First Order Derivative | Second Order Derivative | Magnitude of Average Rotational Speed |
| Dominant Frequency | RMS | Signal Magnitude Area |

Table 4.1: List of features for machine exercise recognition.

### 4.2.4 Exercise Quality Assessment

The final stage of the analytics engine is to assess the quality of the machine exercises performed by exercisers. Assessing the quality of the performed exercise is subjective. A professional trainer usually assesses the quality of a user's exercise based on her body posture, speed of doing the repetition etc. In our case, we have only motion sensor data to assess the quality of the exercise. So we aim to assess the quality of the performed exercises based on motion sensor data. Particularly we aim to compare the user's exercise motion data with trainer's exercise motion in order to find out the similarity with trainer's data. Based on similarity score, we give feedback to the user about how well they are doing during each repetition.

To achieve this goal, we recruited one male professional trainer and one female professional trainer from the fitness center at the university. We collected data of our targeted 12 machine exercises from both trainers and use their data as the guide models for male and female users respectively. Therefore, we assess the quality of participants' exercises by comparing the similarity between participants' exercise data with the guide models. Although the quality of the guide models can be further improved by incorporating data from more trainers, we include only one male and one female trainer to examine the feasibility of our approach.

In this work, we design a motion trajectory based scheme to measure the similarity between each exercise repetition and the guide models. Specifically, the first step of our scheme is to divide the sensor data of each segmented repetition into a sequence of fixed-length tiny windows whose length is much smaller than the duration of the complete repetition itself (the duration of a complete repetition ranges from 3 second to 5 second across different targeted machine exercises. The length of the tiny window we use is 0.5 second). Then we extract a number of features which capture the intrinsic characteristics of each repetition from each tiny window and stack them together to form a local feature vector. As a consequences, each repetition segment is transformed into a sequence of local feature vectors which forms a *motion trajectory* in the feature space. Based on the extracted motion trajectory, we have developed a trajectory comparison algorithm to quantify the similarity between two motion trajectories. It should be emphasized that *the key benefit of the motion trajectory transformation to our quality assessment task in our system is that the motion trajectory provides fine-grained descriptions about where the user's exercise repetition differ from the trainer's guide model*. As such, our platform is capable of providing very concrete feedback to exercisers on how to improve their exercise quality. Below we describe the features we extract from the tiny windows and the details of the trajectory comparison algorithm.

#### 4.2.4.1 Local Feature Extraction

We extract five local features from each tiny window. These features are selected because they capture different aspects of the exercise quality.

- *Average of Movement Intensity (AI):* AI is computed as the average of Motion Intensity (MI) where Motion Intensity (MI) defined as the Euclidean norm of the acceleration vector. AI measures the average strength level of the exercise repetition.

- *Variation of Movement Intensity (VI):* VI is computed as the variation of MI. It measures the strength variation of the exercise repetition.

- *Smoothness of Movement Intensity (SI):* SI is computed as the derivative values of MI. It measures the smoothness of the exercise repetition.

- *Averaged Acceleration Energy (AAE):* AAE calculates the mean value of energy over three accelerometer axes. It measures the total exercise acceleration energy.

- *Averaged Rotation Energy (ARE):* ARE calculates the mean value of energy over three gyroscope axes. It measures the total exercise rotation energy.

#### 4.2.4.2 Motion Trajectory Comparison

The goal of the trajectory comparison is to quantify the similarity between the motion trajectory extracted from an exerciser's repetition and the motion trajectory extracted from the trainer's repetition of the same machine exercise. As such, the exercise quality of the exerciser can be derived. However, one challenge for trajectory comparison is that trajectories from two repetition segments may have different lengths. In this work, we develop the trajectory comparison algorithm based on the *Multidimensional Dynamic Time Warping* (DTW) technique to resolve this issue. DTW is a nonlinear alignment technique for measuring similarity between two signals which have different lengths. In our case, DTW is used to cope with different trajectory lengths. Specifically, let X denote the motion trajectory of the trainer's repetition and Y denote the motion trajectory of the user's repetition whose exercise quality is being assessed:

$$X = x_1, x_2, ..., x_i, ..., x_M \tag{4.1}$$

$$Y = y_1, y_2, ..., y_j, ..., y_N \tag{4.2}$$

where $x_i$ and $y_j$ represent the $i^{th}$ and $j^{th}$ local feature vector in X and Y respectively; and M and N represent the length of X and Y respectively. DTW compensates for the length differences and finds the optimal alignment between X and Y by solving the following dynamic programming (DP) problem:

$$D(i, j) = min\{D(i-1, j-1), D(i-1, j), D(i, j-1)\} + d(i, j) \tag{4.3}$$

where $d(i, j)$ represents the distance function which measures the local difference between local feature vector $x_i$ and $y_j$ in the feature space, and $D(i, j)$ represents the cumulative (global) distance between sub-trajectory $\{x_1, x_2, ..., x_i\}$ and $\{y_1, y_2, ..., y_j\}$. The solution of this DP problem is the cumulative distance between the two trajectories X and Y which sits in D(M, N) and a warp path W of length K

$$W = w_1, w_2, ..., w_k, ..., w_K \tag{4.4}$$

which traces the mapping between X and Y. Since the cumulative distance D(M, N) is dependent on the length of the warp path W, we normalize D(M, N) by dividing it by the warp path length K and use this averaged cumulative distance as the metric to measure the distance between trajectories X and Y as

$$Dist(X, Y) = \frac{D(M, N)}{K} \tag{4.5}$$

In this work, we use the cosine distance as the local distance function defined as

$$d(i,j) = 1 - \frac{{x_i}^T.y^j}{||x_i||.||y_i||} \tag{4.6}$$

Compared to other distance functions, the benefit of using the cosine distance is that d(i, j) is by nature in the range [0, 1]. As a result, the averaged cumulative distance $Dist(X, Y)$ defined in (4.5) is also in the range [0, 1], and thus can be interpreted as the dissimilarity between X and Y in terms of percentile. Therefore, we can define the similarity score in percentile between X and Y as:

$$Sim(X, Y) = 1 - Dist(X, Y) \tag{4.7}$$

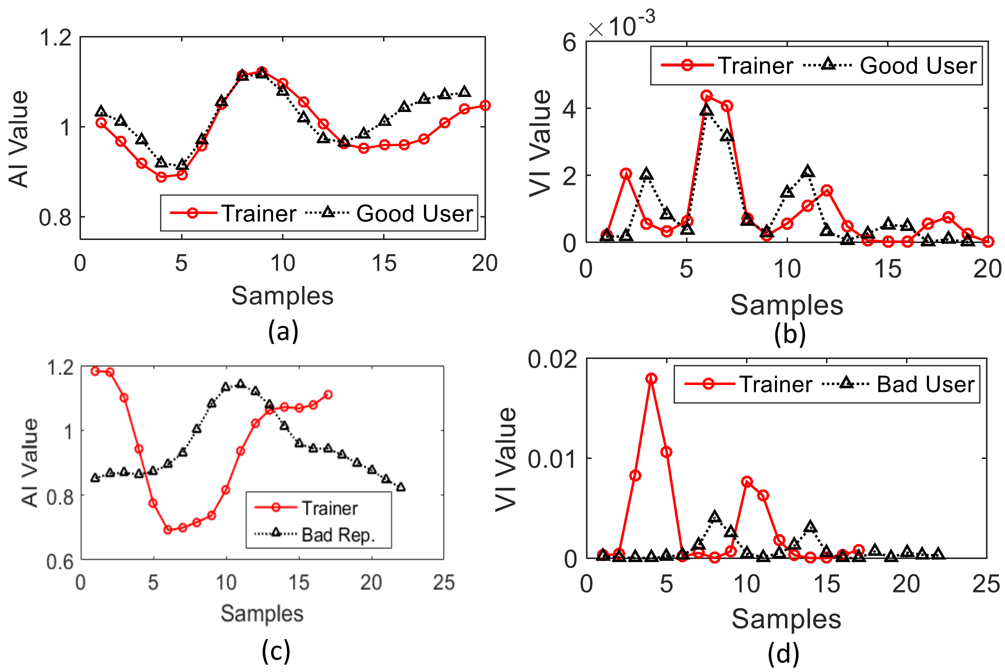



Figure 4.5: (a) and (b): the motion trajectories of one repetition of Leg Extension in feature space of AI (a) and VI (b) with similarity score 98.21%. (c) and (d): the motion trajectories of one repetition of Biceps Curl in feature space of AI (c) and VI (d) with similarity score 16.97%.

Figure 4.5 provides an illustration of our motion trajectory based scheme for exercise quality assessment. In particular, Figure 4.5(a) and Figure 4.5(b) show the fine-grained motion trajectories of one repetition of *Leg Extension* in the feature space of AI and VI respectively. The red trajectories represent the repetition performed by the trainer and the black trajectories represent one high-quality repetition performed by the exerciser. The similarity score between these two repetitions is 98.21%. In comparison, Figure 4.5(c) and Figure 4.5(d) compare a low-quality repetition of *Biceps Curl* with the trainer guide model. The similarity score in this case is 16.97%. As demonstrated, the higher the similarity score is, the better the exercise quality is. We obtain similar results for the other machine exercises.

# Chapter 5

# System Evaluation

## 5.1 Experimental Setup

We recruited 15 participants (11 males and 4 females) who volunteered to help collect data and conduct evaluation experiments. The participants are university students, researchers and stuffs with ages ranging from 22 to 48 years old ($\mu = 27.73$; $\sigma = 6.65$), weights ranging from 42 kg to 85 kg ($\mu = 60.51$; $\sigma = 8.85$), heights ranging from 152cm to 189cm ($\mu = 174$; $\sigma = 6.50$) and experience levels covering novice and intermediate levels. The experiments were conducted at a fitness center on the university campus.

As mentioned earlier, to examine the impact of sensor placement locations on the performance of our system, we attached two sensor tag devices to two different locations on each machine. Figure 5.1 presents an example of the two tag locations on two machines. We envision that in the future, every exercise machine produced by machine manufacturers will have a standardized slot to plug in the tag. Our experiment will help machine manufacturers find the best location for the standardized slots.

During data collection, the participants were instructed to perform the twelve targeted machine exercises by following the short set of instructions on each machine. To capture the intra-subject variability, each participant attended three sessions of data collection. In each session, the participant performed 10 repetitions for each machine exercise. In total, each participant contributed 30 repetitions for each exercise.

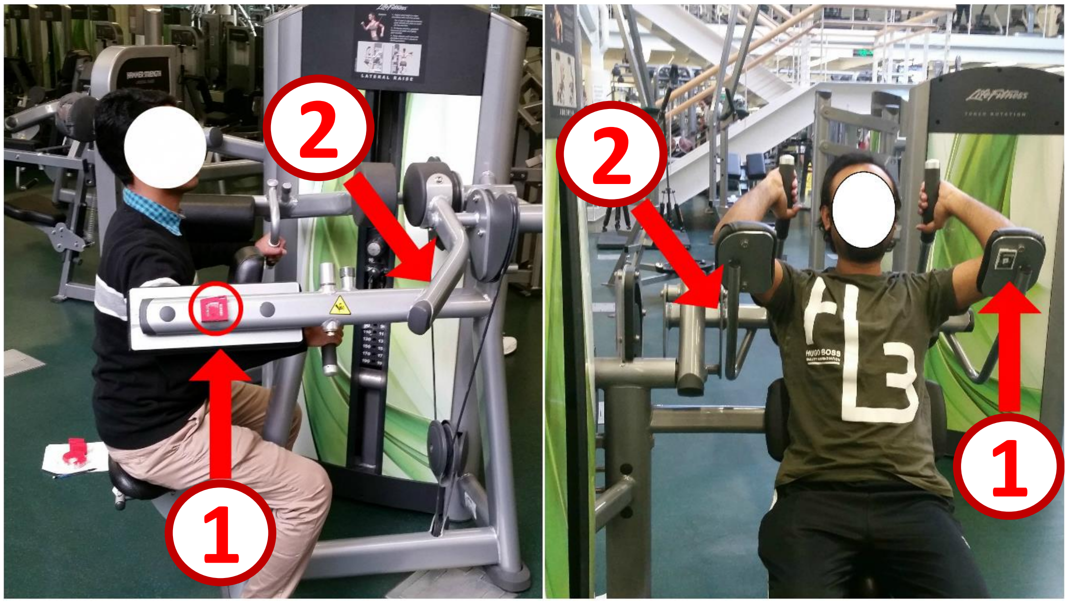



Figure 5.1: The illustration of data collection setup and sensor tag placement using lateral raise machine (left) and seated abs machines (right). Circled 1 and 2 indicate the placement of two sensor tags on each machine.

## 5.2 Performance of Repetition Segmentation and Counting

### 5.2.1 Evaluation Metrics

We evaluate the performance of our repetition segmentation scheme using three metrics:

- *Miss Rate (MSR):* MSR is defined as the proportion of cases that our scheme misses to detect a repetition.
- *Merge Rate (MGR):* MGR is defined as the proportion of the cases that our scheme merges two or more repetitions into one repetition.

- *Fragmentation Rate (FR):* FR is defined as the proportion of the cases that our scheme splits a single repetition into more than one repetitions.

### 5.2.2 Evaluation Results of Repetition Segmentation

We evaluate the performance of our repetition segmentation scheme on our full dataset of 5400 repetitions of 12 exercises from 15 participants. Table 5.1 shows the performance of our repetition segmentation scheme for each type of machine exercises. In terms of MSR, our scheme achieves zero MSR for all machine exercises. In terms of MGR our scheme achieves zero MGR for all machine exercises except E08 *Leg Curl* with the MGR of only 0.11%. Finally, in terms of FR, our scheme achieves zero FR for 6 out of 12 types of machine exercises. Among the other 6 types, the highest FR is only 0.33% for E05 *Triceps Press* and E10 *Row Deltoid*. Taken together, the repetition results indicate that our scheme is able to achieve highly accurate and robust segmentation performance across all types of machine exercises.

### 5.2.3 Evaluation Results of Repetition Counting

Based on the repetitions segmented by our scheme, we achieve the repetition counting accuracy of 97.96% out of a total of 5400 repetitions. Taking a closer look at the counting

| **Exercise** | **E 01** | **E 02** | **E 03** | **E 04** | **E 05** | **E 06** |
|---|---|---|---|---|---|---|
| **MSR (%)** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **MGR (%)** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **FR (%)** | 0.11 | 0.11 | 0.11 | 0.00 | 0.33 | 0.00 |
| **Exercise** | **E 07** | **E 08** | **E 09** | **E 10** | **E 11** | **E 12** |
| **MSR (%)** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **MGR (%)** | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 |
| **FR (%)** | 0.00 | 0.22 | 0.00 | 0.33 | 0.00 | 0.00 |

Table 5.1: Performance of repetition segmentation.

results, we achieve an accuracy of 99.81% for *within 1* scenario (i.e., ±1 count off compared to the ground truth within one session of 10 repetitions) and 100% for *within 2* scenario.

## 5.3 Performance of Exercise Recognition

### 5.3.1 Impact of Sensor Location and Voting

First, we examine the impact of sensor placement location, as well as the session-wise voting mechanism on exercise recognition performance. The goal is to find the best sensor tag deployment location on exercise machines. To achieve this goal, we trained four sets of classifiers 1) from both locations; 2) only from location with larger movement range, 'L'; 3) only from location with smaller range, 'S'; and 4) from the best location combinations. Specifically, the first classier is trained by both locations merged at feature level. Again, the fourth classifier is determined by brute-force searching among 4096 ($2^{12}$) sensor placement combinations. Also, we trained another four sets with aforementioned voting among repetitions in the same session. The evaluated results are listed in Table 5.2 by using leave-one-subject-out cross validation strategy.

The maximum accuracy of 99.08% is achieved from the best location combination with the session-wise voting. Different location combination helps to distinguish between exercises with similar (such as- Seated Abs, Leg Extension and Lateral Raise have very similar movement pattern) movement pattern using different range of motion from different locaiotn of the machines. Our further investigation for the latency of the voting mechanism confirmed that the best accuracy can be achieved quickly, after 1.22 repetitions on average (SD=0.92). By comparing the results of the system in a lateral fashion, we understand that the system performs better when it uses session-wise voting, which motivates us to use this technique to increase the performance. Second, by comparing the results in a longitudinal fashion, we find that best location classifier outperforms the rest three classifier with around 10% point difference. This implies that our study of impact of sensor location can guide the manufacturer and user to find and mark the best locations.

| **Datasets** | **Without session-wise voting** | | | | **With session-wise voting** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Precision** | **Recall** | **Accuracy** | **F-measure** | **Precision** | **Recall** | **Accuracy** | **F-measure** |
| All locations | 0.8325 | 0.8266 | 82.67% | 0.8295 | 0.8970 | 0.8887 | 88.88% | 0.8930 |
| Location 'L' only | 0.8102 | 0.8081 | 80.82% | 0.8091 | 0.8854 | 0.8809 | 88.10% | 0.8831 |
| Location 'S' only | 0.8114 | 0.8121 | 81.22% | 0.8117 | 0.8829 | 0.8815 | 88.17% | 0.8822 |
| Best Locations | 0.9432 | 0.9430 | 94.30% | 0.9431 | 0.9911 | 0.9907 | 99.08% | 0.9909 |

Table 5.2: Overall precision, recall, accuracy and F-measure for different sensor locations and voting option.

### 5.3.2 Performance of Subject Independent Model

Next, we examine the exercise type recognition performance of the subject independent model. Figure 5.2 shows that the average recognition accuracy across all the machine exercises using leave-one-subject-out cross validation. As shown, 11 out of 15 subjects achieve 100% accuracy while the other 4 subjects achieve accuracy of 97.24%, 97.22%, 97.22%, and 94.45%, respectively. In addition, the confusion matrix in Table 5.3 provides a detailed look at the results in terms of types of machine exercises. As shown, 10 out of 12 machine exercises achieve 100% precision and recall. Taken together, our results indicate that JARVIS can achieve high recognition performance across all the targeted machine exercises in a subject independent manner.

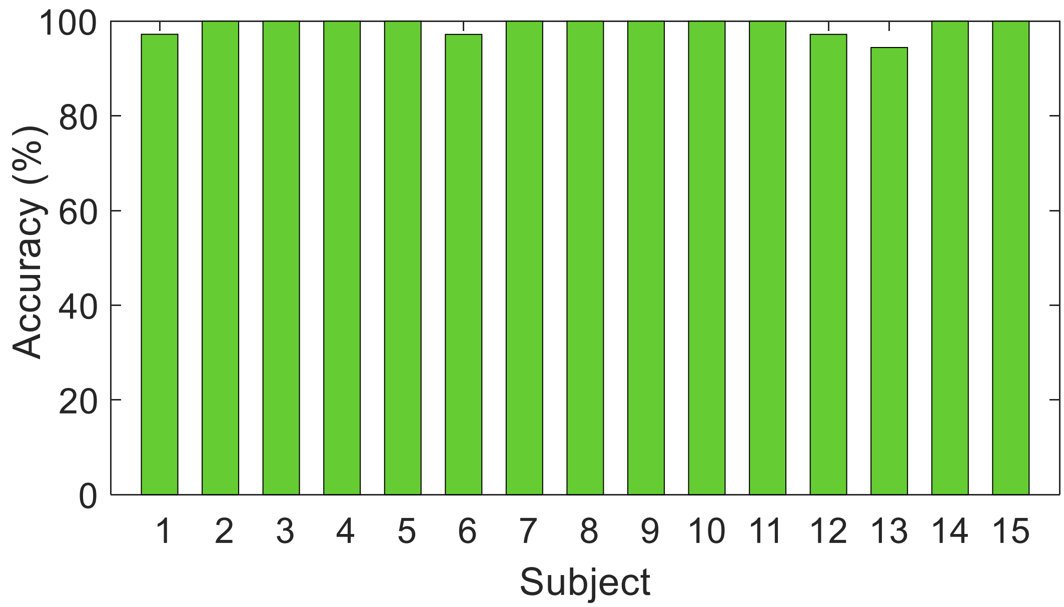



Figure 5.2: Performance of subject independent model.

### 5.3.3 Impact of Number of Features

Finally, we examine the impact of the number of features on the performance of exercise type recognition. Figure 5.3 shows the average recognition accuracy related to the number of features selected by the sequential floating forward selection (SFFS) algorithm using leave-one-subject-out cross validation. We use linear SVM as our classifier for feature selection. As shown, the recognition accuracy increases in general as the number of features increases.

| | **E01** | **E02** | **E03** | **E04** | **E05** | **E06** | **E07** | **E08** | **E09** | **E10** | **E11** | **E12** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **E01** | 451 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E02** | 0 | 451 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E03** | 0 | 0 | 451 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E04** | 0 | 0 | 0 | 450 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E05** | 0 | 0 | 0 | 0 | 452 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E06** | 0 | 0 | 0 | 0 | 0 | 450 | 0 | 0 | 0 | 0 | 0 | 0 |
| **E07** | 0 | 0 | 0 | 0 | 0 | 20 | 420 | 0 | 10 | 0 | 0 | 0 |
| **E08** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 451 | 0 | 0 | 0 | 0 |
| **E09** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 450 | 0 | 0 | 0 |
| **E10** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 430 | 0 | 20 |
| **E11** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 450 | 0 |
| **E12** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 451 |

Table 5.3: Confusion matrix of exercise type recognition.

The accuracy reaches the highest when using 23 features out of a total of 28 features. The most significant 23 features are: Mean, Median, Standard Deviation, Variance, Skewness, Kurtosis, Energy, Interquartile Range, Spectral Entropy, First Order Derivative, Second Order Derivative, Magnitude of Average Rotational Speed, RMS. We also observe that the recognition accuracy already reaches 98% when only using 12 features. This result indicates that by using a small number of features, JARVIS could still achieve high recognition performance.

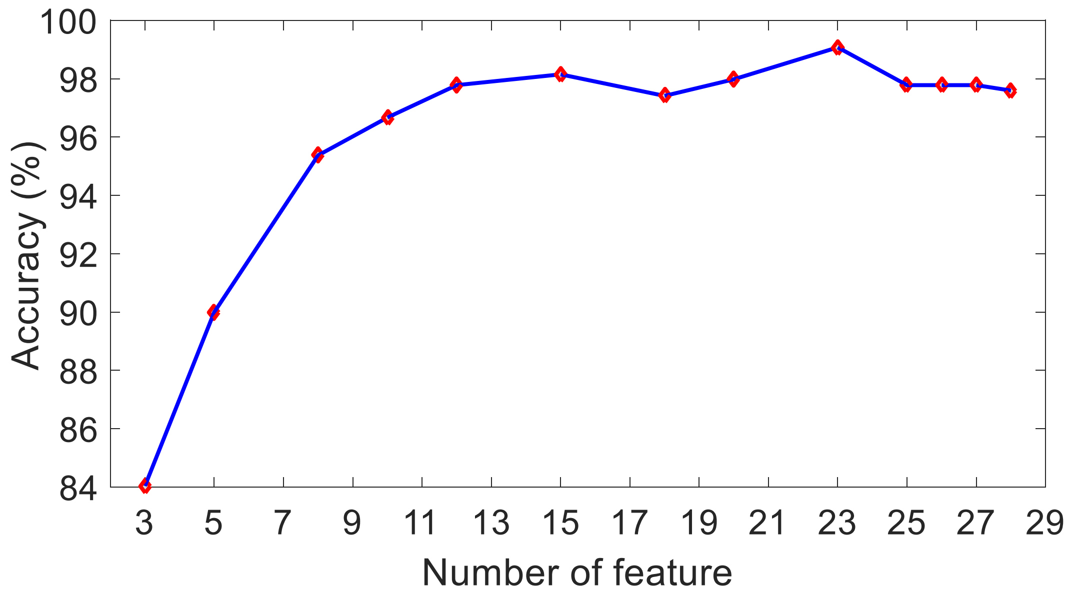


Figure 5.3: Accuracy changes over number of features.

## 5.4 Validity of Quality Assessment

To examine the validity of our exercise quality assessment scheme, we use the two trainers' data. Due to the strength difference between male and female trainer, we compare the male/female trainer's exercise data to his/her own data. Specifically, we asked each trainer to perform the 12 targeted machine exercises. Each machine exercise was performed 10 repetitions per session for three sessions. Given these data, we calculate the similarity scores for all pairs of repetitions within each machine exercise for male trainer and female trainer separately. We expect the similarity scores to be high because both of them are professional trainers who perform consistent high-quality machine exercises. Figure 5.4 illustrates the average similarity scores related to types of machine exercises for both main trainer and female trainer. As shown, average similarity scores are high across all the machine exercises for both trainers. In particular, the lowest average similarity score is 90.80% and 88.25% for male trainer and female trainer respectively. This result indicates that our quality assessment scheme is able to reliably assess the exercise quality across different types of machine exercises for both male and female.

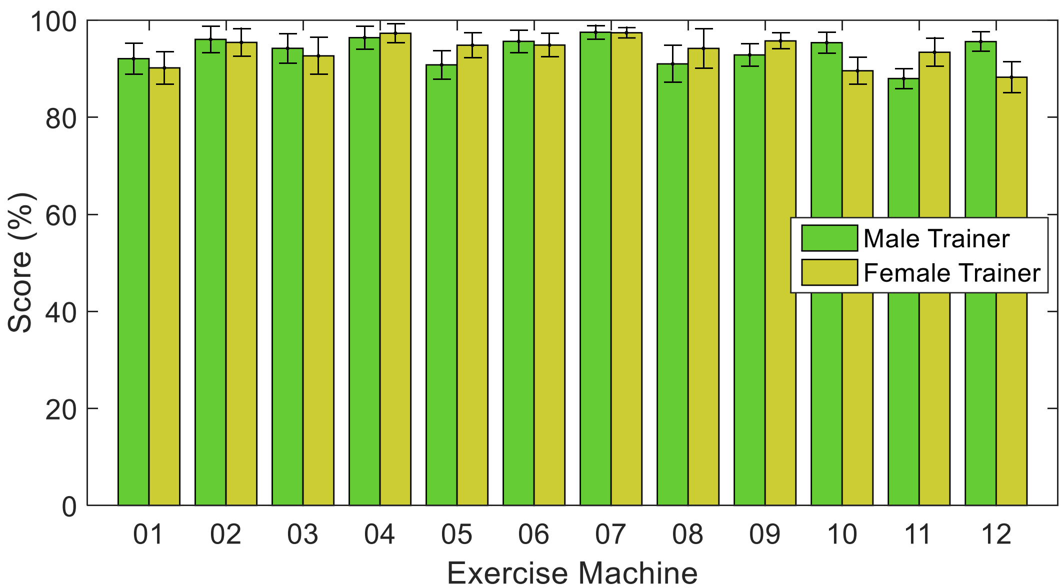



Figure 5.4: Validity of exercise quality assessment.

## 5.5 System Performance

In this section, we report system performance profiles of JARVIS, including processing time, energy consumption, and temperature changes. For the tests, we employ Samsung Galaxy S6 smartphone, having quad-core 2.1 GHz Cortex-A57 processor and 3 GB ram, and run-

| Tasks | Time required (per rep.) (ms) |
|---|---|
| Pre-processing | 0.012 |
| Segmenting Repetition | 0.018 |
| Feature Extraction | 0.006 |
| Classification (using libSVM) | 17.5 |
| Quality Assessment | 0.457 |

Table 5.4: Processing time statistics.

ning Android OS v5.0.2 (Lollipop), mounted to Samsung Gear VR Headset. Also, we use CC2650STK SensorTag as a sensing device.

### 5.5.1 Processing Time

To examine the computational demand of our exercise analytics engine, we measure the average processing time consumed by each components of the engine. The computational pipeline consists of 5 core steps shown in Table 5.4. For the classification component, we employ the Android implementation of libSVM [2]. We run the 5 set of processing trials using 100 repetitions of different exercises data from different participants. We calculated the average processing time using the results from these 5 set trails. Table 5.4 shows that our analytics engine can run fast with minimum computational overhead. Among all analytics engine components, the exercise type recognition takes the most significant amount of time among the analytics engine components.

### 5.5.2 Energy Consumption

#### 5.5.2.1 VR HMD Energy Consumption

Due to the immobility of our setup as shown in Figure 5.5, it is impractical to measure the real-time energy consumption of JARVIS in gym setting [19]. Instead, in lab setting, we let the system still receive data packets from the sensor tag, and fed the exercise analytics engine the previously-collected exercise data. We pre-loaded all records in memory to eliminate additional flash storage access during measurement. We use Monsoon Power Monitor device [4] to measure energy consumption of computation and communication tasks. To exactly measure the energy consumed for our system, we turned off irrelevent services and components including GPS, WiFi and cellular services. We also shut down all other applications. We profile the energy consumption per computational component. We measured the energy consumption 5 times and we take the average value of the trials. Table 5.5 shows the breakdown of energy consumption for each component. Our calculation indicates that JARVIS can run 2.85 hours on Galaxy S6 smartphone, which has 2550 mAh Li-Ion battery. For VR Synthesis Frontend, we tested three versions providing different graphics rendering quality levels, which will be further described in the device temperature measurement. The three settings showed up to 8.45% difference in terms of energy consumption.

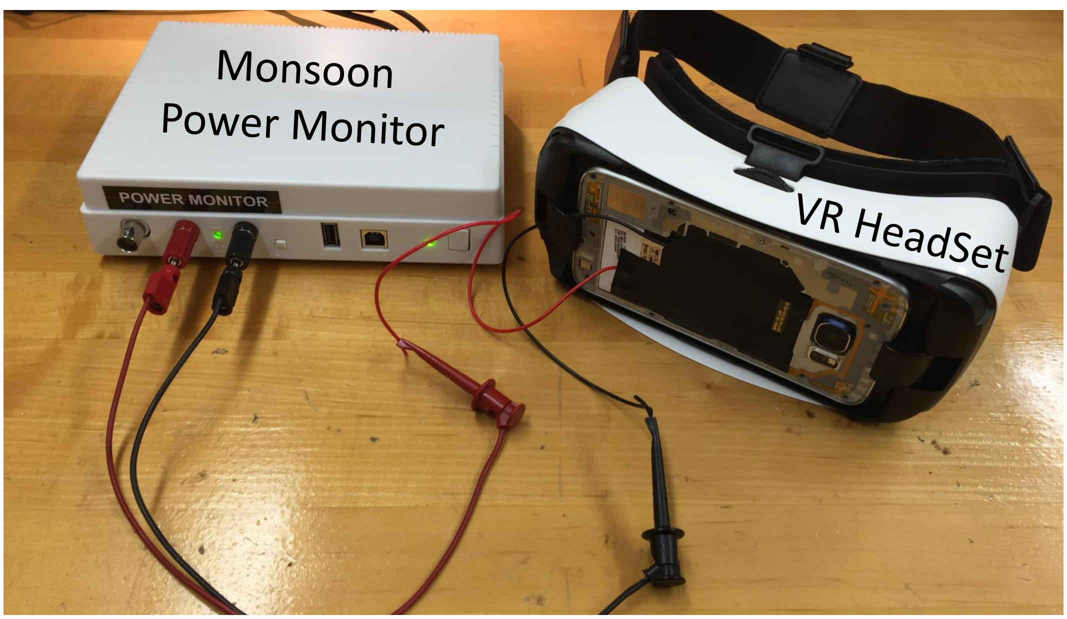



Figure 5.5: Energy consumption measurement setting with Samsung Galaxy S7, Gear VR and Monsoon Power Monitor.

| Device | Component | | Current (mA) | Power (mW) |
|---|---|---|---|---|
| Phone | BLE Communication | | 61.46 | 245.7 |
| | Processing Backend | | 17.65 | 70.57 |
| | VR Frontend | Low Quality | 800.3 | 3198 |
| | | Mixed Quality | 815.6 | 3259 |
| | | High Quality | 867.9 | 3468 |
| SensorTag | 10Hz Data Transmission | | 3.967 | 11.90 |

Table 5.5: Energy consumption statistics

#### 5.5.2.2 SensorTag Energy Consumption

We measured and calculated the energy consumption and estimated battery lifetime of the TI CC2650STK SensorTag, powered by a 3V, 240mAh lithium coin cell battery. We powered the sensor tag by Monsoon power monitor and let it transmit motion sensor data to the smartphone at 10HZ. Table 5.5 shows the details. From our calculation, we find that the battery can last 60.5 hours.

### 5.5.3 Device Temperature and Visual Quality

We measured temperature changes of the VR HMD device over time, with three different implementation settings in rendering quality: (1) High-Quality, (2) Mixed-Quality, and (3) Low-Quality (see Figure 4.2 for example screenshots). High-Quality setting has 155.8k triangles to render on average in the game scene, with a bumped specular shader with lightmap. Low-Quality has 13.8k triangles in the game scene, with a standard shader for the virtual body. Mixed-Quality has 30.1k triangles, with increased number of triangles and a bumped specular shader with lightmap on target muscle groups (e.g., abdominals). We employed the log features using Android Debug Bridge (adb) shall to track temperature changes, and collected CPU and Ambient temperature values. Our indoor testbed had room temperature of about 25 degree Celcius. We ran the VR application for 30 minutes.

As shown in Figure 5.6, the CPU temperature measurement result indicated that Mixed-

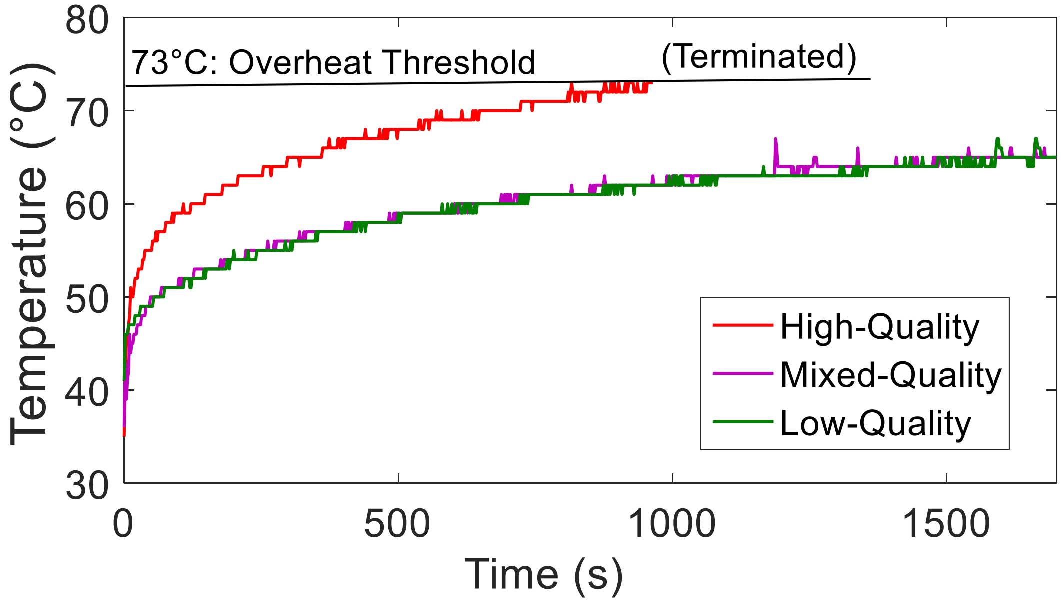



Figure 5.6: CPU temperature changes over time. High-Quality overheated the phone after 15 min.

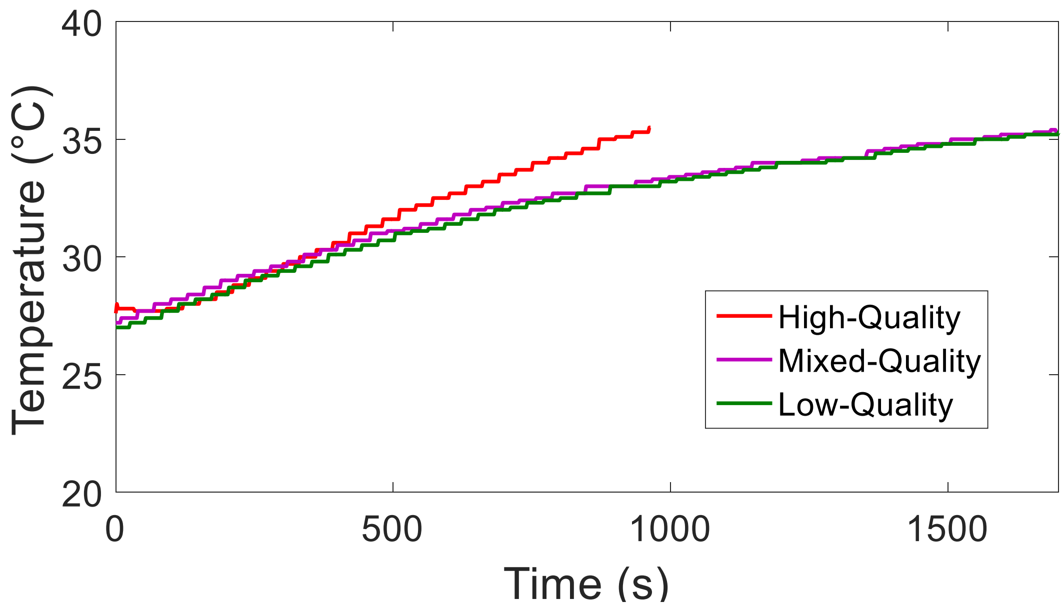



Figure 5.7: Ambient temperature changes over time.

and Low-Quality settings allowed the VR HMD hardware to run the application for at least 30 minutes. Temperature levels for the two settings were under 70 degree after 30 minutes of operation. However, High-Quality setting overheated the CPU up to 73 degree in 15 minutes, and the application was terminated for device safety. Ambient temperatures (in

Figure 5.7) in Mixed- and Low-Quality reached to 35 degree after 30 minutes, whereas one in High-Quality did in 15 minutes.

We carefully suggest that it is important to help developers estimate the effect of visual quality to temperature changes, and believe that our experiment result can be one helpful case for better temperature estimation.

# Chapter 6

# Case Study with JARVIS

In this case study, we evaluate the efficacy of our exercise tracking platform app with questionnaire survey and EMG signal analysis, as well as interview for its overall user experience.

We set the seated abs as a target exercise, due to its importance for health benefits. Usually abdominal training is considered to build the *rectus* muscles around bellybutton, however, abdominal muscles are very sensitively connected to each other, and to lumbar spine as well. It is known that abdominal muscle training is beneficial for lumbar stability, however, without correct guidance it primarily activate rectus abdominis, while oblique abdominis muscles (highlighted muscle group in Figure 4.1) are considered to be more import contributors to lumbar stability [20]. One effective intervention for balanced abdominal training is verbal instructions, which showed statistically significant differences in terms of abdominis muscle activation measured by surface electromyography (sEMG) [14]. The goal of our case study is to see whether our muscle highlighting feature can help people to concentrate on a particular muscle and whether it can help people to activate that particular muscle. In this case study we capture the muscle activity data using a surface electromyography (sEMG) sensing device to capture a particular muscle activity.

## 6.1 Study Methods

### 6.1.1 Participants and Instrumentation

We recruited 10 students that are different from the participants for earlier data collection, through department email distribution lists, as well as on-site recruitment at a fitness center in the university campus. 5 students were hired through each way of recruitment, respectively. All participants had prior experience with the seated abs machine.

EMG data were collected using Trigno Wireless sEMG System [1], which consists of one base station and multiple wireless sEMG sensors. Each sEMG sensor has signal bandwidth of 20-450 Hz, transmission range of 40 meters and sampling rate of 4000 samples/sec with 16-bit resolution. The base station is capable of streaming data to a manufacturer-provided analysis software over USB wired connection. The transmitted data was saved into computer storage for analysis.

### 6.1.2 Study Design and Analysis

Following the abdominal exercise study [14], 4 sEMG sensors were placed on the left anterior abdominal wall, at upper/lower rectus abdominis (URA/LRA), and external/internal oblique abdominis (EOA/IOA). (See Figure 6.1)

Participants were instructed in two conditions: one condition with our exercise tracking application emphasizing activation of oblique abdominis, and one controlled condition without the application. Prior to the condition with the exercise tracking platform, participants were told about the basic user interfaces of the system, and asked to take a look at the VR space with their virtual body. In the controlled condition, the participants are asked to perform the exercise naturally, as if they usually would at gyms. For each condition,

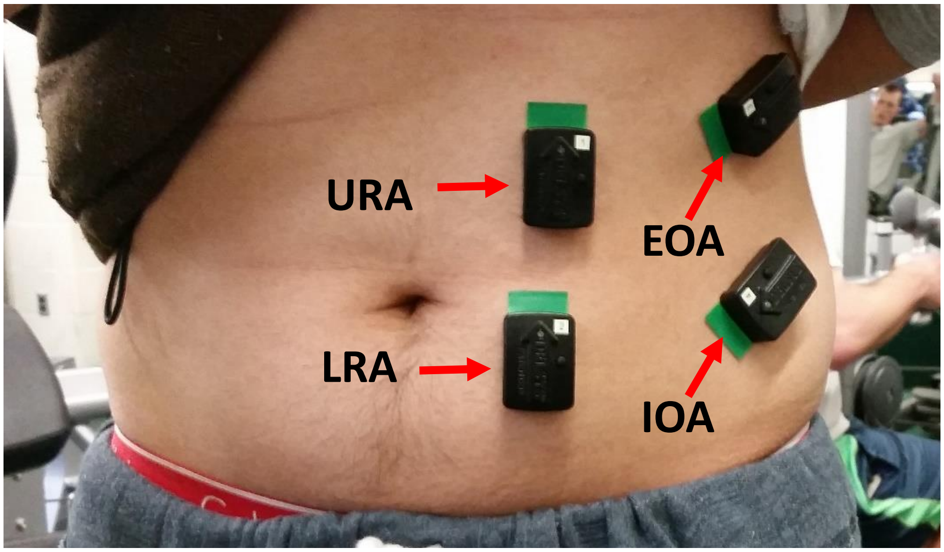



Figure 6.1: Illustration of sEMG sensor placement at upper/lower rectus abdominis (URA/LRA), and external/internal oblique abdominis (EOA/IOA).

the participants performed two sets of seated abs exercise with ten repetitions, respectively. sEMG activity was recorded from the beginning to the end of each set. The participants were allowed to rest for up to two minutes between sets. The sequence of the conditions are balanced across the participants.

After each condition, the participants were asked to respond to 19 selected and adjusted questions from intrinsic motivation inventory (IMI) [26, 29], which is a widely used and extensively validated set of questionnaire to evaluate user experiences. Among five major categories of IMI, we chose three categories related to the application and exercise context: interest/enjoyment, perceived competence, and value/usefulness, having 7, 5 and 7 questions respectively. The responses were collected in 7-point Likert scale and then averaged in each category for statistical analyses.

A semi-structured interview followed after the end of study. We asked the participants about their overall exercise experiences, including their perception about the VR application and its interface, the efficacy of the application, as well as their suggestions in design. All interview data were transcribed and analyzed by open coding [31] and axial coding [12] to discover common themes and patterns.

## 6.2 Results

### 6.2.1 EMG activity

Figure 6.2 shows representative sEMG data from individual trials performed under conditions with and without VR respectively. In this figure, it is evident that sEMG amplitudes of the abdominis muscles change in response to repetitions. These changes in muscle activity are in concert with the changes in group mean sEMG values shown in Figure 6.3.

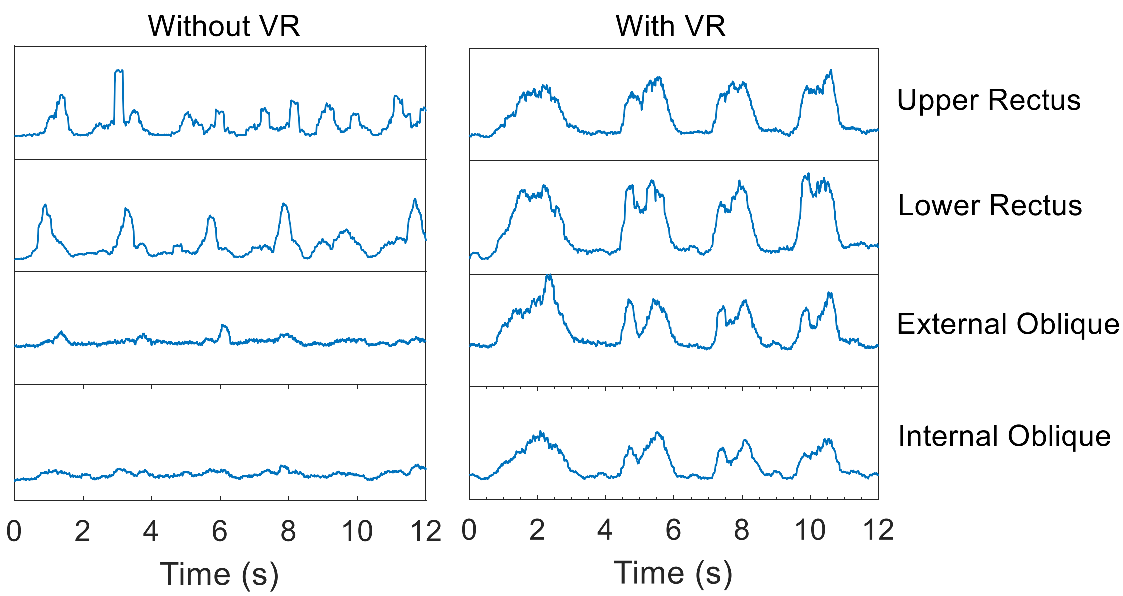



Figure 6.2: Without VR vs. With VR in terms of the root-mean-square (RMS) values of surface EMG signals.

Figure 6.3 provides summary data from the two conditions, with group means and standard deviations of the normalized EMG activity of each of the 4 muscle groups. There was a significant effect of VR application for the two oblique muscle group ($p < 0.01$ and $p < 0.05$ for extenal and internal oblique abdominis respectively), which shows efficacy of the VR application in activating target muscle groups, compared to the previous study in the literature of orthopaedic and sports physical therapy [14].

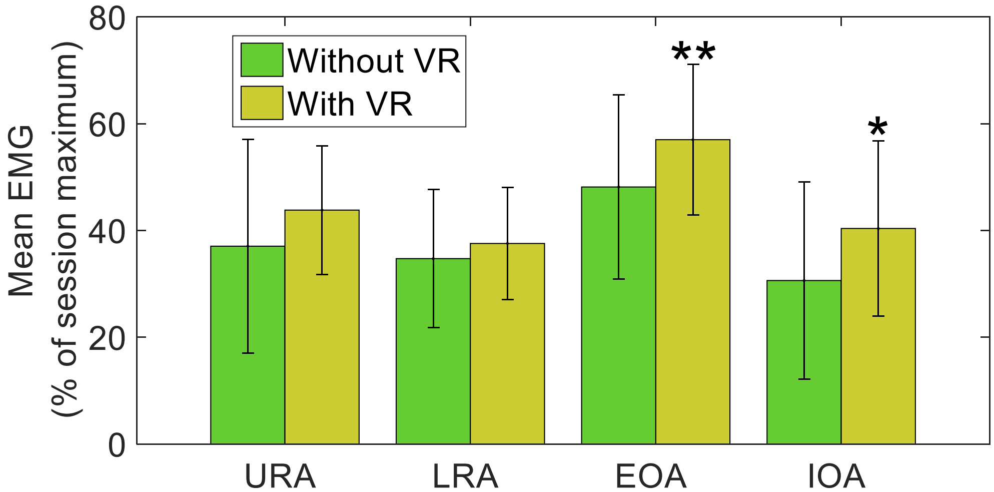



Figure 6.3: Mean normalized electormyography values for four muscle groups. With-VR significantly different from Without-VR at oblique abdominis (**$p < .01$; *$p < .05$).

### 6.2.2 User experience: Questionnaire survey

Figure 6.4 shows summary data from the responses of IMI questionnaire survey. For all three categories, there was a significant effect of VR application ($p < 0.05$). In-depth analysis revealed that the students recruited through the mailing list, which had less experiences with strength training machines, showed greater differences in all three categories, than ones hired on-site at the fitness center. This gives a hint for relationship between users' experience level and engagement. That is, the VR application has a potential in motivating novice or less-experienced exercisers to engage in strength training, implying an opportunity of broader impact to public health.

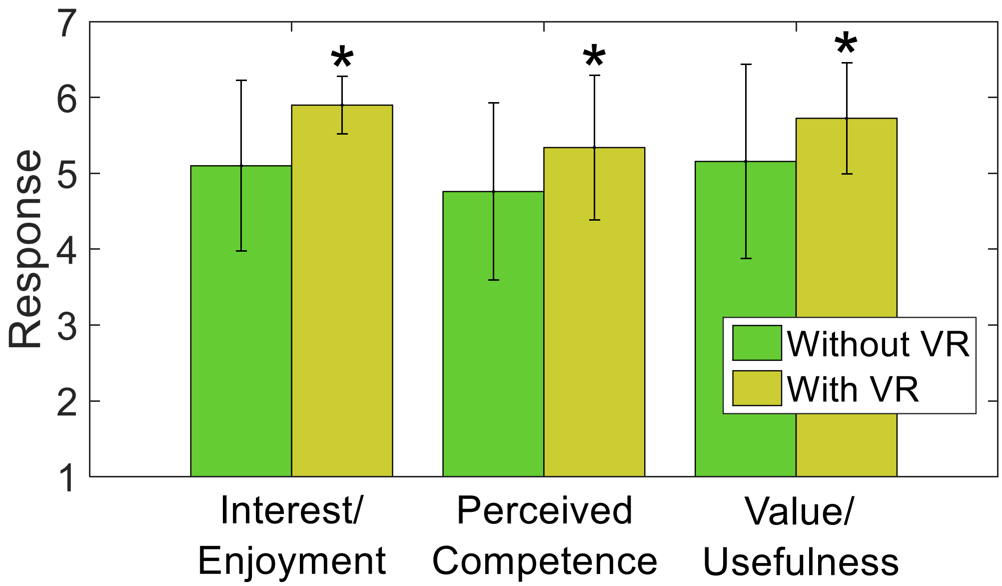



Figure 6.4: IMI survey results for three categories. With-VR significantly different from Without-VR (*$p < .05$ for all three categories).

### 6.2.3 User experience: Interview analysis

The participants shared their perceptions about JARVIS application including its benefits and usefulness, as well as their suggestions to provide *"sweet"* (P7) and more enjoyable exercise experiences. In overall, they enjoyed and valued the application and its features, which is consistent to the earlier IMI survey results.

#### 6.2.3.1 Virtual Self and its Movement

The participants liked JARVIS which enabled them to *"see myself moving"* (P8). Seeing their virtually represented body in VR helped them in doing *"the exercise a lot more correctly"* (P6) and in better *"focusing on (my) body"* (P6). Some participants talked about the discrepancy between real self and virtual self, in terms of their appearance: *"[The man in VR] was not exactly the same (to me), but the motion I was like the same with him."* (P4) Interestingly, the discrepancy showed a positive potential for user motivation: *"looking at the muscle man really attracted me to follow him."* (P1) In overall, 'virtual self' contributed to *"more enjoyable"* (P8) exercise experience, consistent to the earlier IMI-enjoyment survey results.

#### 6.2.3.2 Benefits of Real-time HUD

The participants liked the real-time informative HUD user interface, providing *"consistent feedback about time and quality"* (P4): *"You could definitely better simply because it tells you the efficiency of the workout you are doing."* (P7) They recalled how hard it was to *"count the repetitions while focusing on the exercise"* (P2), and talked that *"(easily) seeing how many repetitions I have"* (P3) was very helpful. Also, with the duration and quality information they could *"run clean repetitions"* (P9), and found the app played a role as a *"binder to improve (their) exercise quality"* (P3).

#### 6.2.3.3 Benefits of Muscle Highlighter

In providing visual guide for required muscle groups to activate, the participants found that the muscle highlighter feature was helpful, otherwise they *"would never know how to do it properly"* (P8): *"What the machine is targeting is the muscle, so highlighting really helped me to perform well (and) instructed me."* (P1) It naturally instructed the participants *"how to focus (their) abdominal part"* (P8), and assisted them to better *"conceptualize the way of exercise"* (P10), contributing to the efficacy of the app validated earlier in muscle activation analysis.

#### 6.2.3.4 Points of Improvements and Design Suggestions

The participants suggested interesting points of improvements for JARVIS application. One participant talked about a soft competition with another pacemaker avatar or friends online in the VR screen, related with Köhler effect, a well-known theory about benefits of competition in the literature of sports science [24, 16]: *"I would make a goal to reach to keep it up with a guy in the VR."* (P8) They also wanted to try the application for different machines which may be also beneficial: *[it would be helpful for] other popular exercise I know like chest press. That will be very helpful because a lot of people do need better form in that. So those compound exercise would definitely help people."* (P10).

# Chapter 7

# Conclusion and Future Works

In this thesis, we presented the design, implementation and evaluation of JARVIS, a first-of-its-kind sensing system, based on a miniature IoT sensing device combined with a mobile VR headset to enable immersive and interactive gym exercise experience. JARVIS leverages immersive VR technology to provide users with an interactive exercise analytics for machine exercises in real time. The realization of such aim requires the VR headset to retrieve multidimensional information (i.e. repetition count, exercise type, duration of each repetition, quality score etc.) of the machine exercise performed by the exerciser in real-time. JARVIS achieves this by attaching miniature IoT sensing devices on gym machines to capture machine exercises as well as developing online analytics algorithms and VR performance optimization techniques to track exercise progress and assess exercise quality in real time. By converting the exercise progress and quality information to VR inputs, JARVIS creates a truly immersive gym exercise experience with a virtual avatar to guide machine exercises in a highly interactive manner and providing the exercise analytics to the user in real time in a convenient way using VR technology. Through extensive experiments, we have demonstrated that JARVIS could provide real-time machine exercise analytics. Our real-world deployment study indicates that JARVIS could provide an engaging and effective exercise assistance to exercisers for strength training.

Our future work includes evaluating the efficacy of JARVIS with broader population, designing more entertainment-oriented applications and games for gym settings, and extending JARVIS as a full-pledged gym context monitoring and fitness management platform. We envision that JARVIS will eventually redefine gym activities as fully immersive VR experiences.

# BIBLIOGRAPHY

# BIBLIOGRAPHY


[1] Delsys inc. In *http://www.delsys.com/*.

[2] libsvm. In *http://www.csie.ntu.edu.tw/ cjlin/libsvm/*.

[3] Microsoft hololens. In *http://www.microsoft.com/microsoft-hololens*.

[4] Monsoon. In *https://www.msoon.com/LabEquipment/PowerMonitor/*.

[5] Oculus rift. In *https://www.oculus.com/en-us/*.

[6] Samsung gear vr. In *http://samsung.com/galaxy/wearables/gear-vr*.

[7] Sensortag from texas instruments. In *http://www.ti.com/*.

[8] K. Adams, E. Cafarelli, A. Gary, C. Dooly, S. Matthew, S. J. Fleck, A. C. Fry, J. R. Hoffman, R. U. Newton, J. Potteiger, M. H. Stone, N. A. Ratamess, and T. Triplett-mcbride. Progression models in resistance training for healthy adults, 2009.

[9] K.-H. Chang, M. Y. Chen, and J. Canny. *Tracking free-weight exercises*. Springer, 2007.

[10] Y. Degtyarev, E. Cuervo, and D. Chu. Demo: Irides: Attaining quality, responsiveness and mobility for virtual reality head-mounted displays. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, pages 443–443. ACM, 2015.

[11] 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*, pages 141–154. ACM, 2015.

[12] K. Holtzblatt and S. Jones. Contextual inquiry: A participatory technique for system design. *Participatory design: Principles and practices*, pages 177–210, 1993.

[13] A. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(2):153–158, 1997.

[14] G. M. Karst and G. M. Willett. Effects of specific exercise instructions on abdominal muscle activity during trunk curl exercises. *The Journal of orthopaedic and sports physical therapy*, 34(February 2001):4–12, 2004.

[15] L. Karvitz. Exercise Motivation: What Starts and Keeps People Exercising. *University of New Mexico*, 2011.

[16] O. Köhler. Über den Gruppencrirkungsgrad der menschlichen Körperarbeit und die Bedingung optimaler kollektivkraftreaktion [Human physical performance in groups and conditions for optimal collective performance]. *Industrielle Psychotechnik*, 4:209–226., 1927.

[17] P. F. LaChance and T. Hortobagyi. Influence of Cadence on Muscular Performance During Push-up and Pull-up Exercise, 1994.

[18] C. L. Lox, K. A. M. Ginis, and S. J. Petruzzello. *The psychology of exercise: Integrating theory and practice*. 2006.

[19] A. Maiti and G. Challen. The missing numerator: Toward a value measure for smartphone apps. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, pages 99–104. ACM, 2015.

[20] M. I. Miller and J. M. Medeiros. Recruitment of internal oblique and transversus abdominis muscles during the eccentric phase of the curl-up exercise. *Physical therapy*, 67(8):1213–1217, 1987.

[21] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner. Recofit: using a wearable sensor to find, recognize, and count repetitive exercises. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3225–3234. ACM, 2014.

[22] B. J. Mortazavi, M. Pourhomayoun, G. Alsheikh, N. Alshurafa, S. I. Lee, and M. Sarrafzadeh. Determining the single best axis for exercise repetition recognition and counting on smartwatches. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*, pages 33–38. IEEE, 2014.

[23] M. Muehlbauer, G. Bahle, and P. Lukowicz. What can an arm holster worn smart phone do for activity recognition? In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pages 79–82. IEEE, 2011.

[24] K. a. Osborn, B. C. Irwin, N. J. Skogsberg, and D. L. Feltz. The Köhler effect: Motivation gains and losses in real sports groups. *Sport, Exercise, and Performance Psychology*, 1(4):242–253, 2012.

[25] T. Park, I. Hwang, U. Lee, S. I. Lee, C. Yoo, Y. Lee, H. Jang, S. P. Choe, S. Park, and J. Song. Exerlink: enabling pervasive social exergames with heterogeneous exercise devices. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 15–28. ACM, 2012.

[26] R. W. Plant and R. M. Ryan. Intrinsic motivation and the effects of self consciousness, self-awareness, and ego involvement: an investigation of internally controlling styles. *Journal of Personality*, 53(3):435–449, 1985.

[27] A. Rizzo, J. Difede, B. O. Rothbaum, J. M. Daughtry, and G. Reger. Virtual reality as a tool for delivering ptsd exposure therapy. *Post-Traumatic Stress Disorder: Future Directions in Prevention, Diagnosis, and Treatment*, 2013.

[28] C. Rojek. *Leisure and culture*. Macmillan Press, 2000.

[29] R. M. Ryan, J. P. Connell, and R. W. Plant. Emotions in nondirected text learning. *Learning and Individual Differences*, 2(1):1–17, 1990.

[30] B. J. Schoenfeld and B. Contreras. Attentional Focus for Maximizing Muscle Development. *Strength and Conditioning Journal*, 38(1):27–29, 2016.

[31] A. Strauss and J. Corbin. *Basics of qualitative research: grounded theory procedure and techniques*, volume 13. Sage, London, 1990.

[32] T. Wakahara, A. Fukutani, Y. Kawakami, and T. Yanai. Nonuniform muscle hypertrophy: Its relation to muscle activation in training session. *Medicine and Science in Sports and Exercise*, 45(11):2158–2165, 2013.

[33] T. Wakahara, N. Miyamoto, N. Sugisaki, K. Murata, H. Kanehisa, Y. Kawakami, T. Fukunaga, and T. Yanai. Association between regional differences in muscle activation in one session of resistance exercise and in muscle hypertrophy after resistance training. *European Journal of Applied Physiology*, 112(4):1569–1576, 2012.

[34] M. Zhang and A. A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 92–98. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011.