Auto Training Routine Recording System

Haesung Oh

I. Introduction

1. Necessity and Purpose

The development of skeletal muscle mass in weight training is the ultimate goal. Weight-training tears muscles and muscles grow by filling the torn spots. However, if the muscle always takes the same stimulus, the muscles adapt to the stimulus, which slows down muscle growth. Thus, 'Gradual Overload' is considered a vital exercise method in weight training.

Gradual overload is a method of gradually increasing one of the sets, the repeats, and the weight for each exercise. In order to practice this, it is essential to record the daily exercises. However, weight training beginners may need to become more familiar with the routine recording itself. Even for middle and high-class people, there is the hassle of recording it directly on paper or a smartphone notepad whenever they finish every single set.

This paper suggests a system that automatically progresses routine recordings of squats, deadlifts, and bench presses, called three major exercises in weight training, based on a computer vision with only a camera and a computer.

2. Current Problems

Smartwatches are typical hardware that helps with training. However, this smartwatch mainly checks the heart rate to assist with exercise, and crucially, there is a barrier that the user must preset what exercise he or she performs.

Another piece of hardware is the 'Nintendo Switch.' In this game console, a game called <Ring Fit> allows users to play games with various types of exercise. Pressure and gyro sensors can detect force and body movements, but there is a passive aspect of having to do the exercises required by the game.

Figure 1 Smartwatches

Figure 2 Nintendo Switch and <Ring Fit>

In software, there are many smartphone training applications on the market. However, a common feature of these applications is that users must set their routines in advance, including the type and weight of exercise, and directly touch and record the number of repeats during training.

There are also training aids that encompass software and hardware. For example, smart gym, which has been increasing recently, installs cameras for each exercise device and analyzes users' movements with specific programs. However, this also has the disadvantage of having to select and train in advance what kind of exercise the user will do.

In the end, there is a demand for training assistance tools that completely automate to record the type of exercise, weight, number of sets, and repeats.

II. Research Details

1. Development environment

1.1 OpenCV, OpenPose

Since it is necessary to estimate a person's posture with computer vision, the OpenPose algorithm, which is open-source, is used. OpenPose is an OpenCV-based algorithm that tracks body movements with 18 key points.

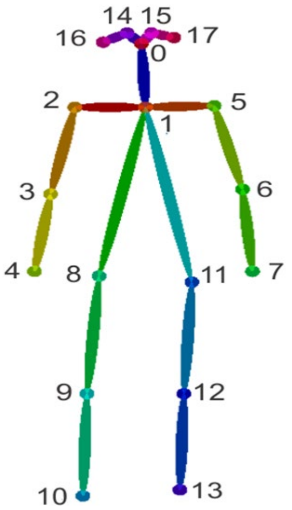

Figure 3 Key points of OpenPose

1.2 Keypoints

Figure 3 shows the 18 key points of OpenPose. These key points, which consist mainly of joints, are obtained and stored in real-time every frame in the video.

1.3 API

There is a Point2f class in OpenPose; the float type x and y coordinates exist as a public member. Eighteen Point2f objects form a vector and update on every video frame. The system uses this 36-dimensional vector to represent a person's posture. Figure 4 shows the approximate structure of the vector.

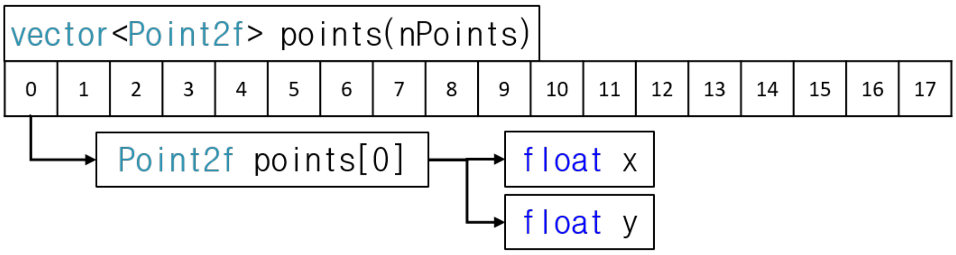

Figure 4 Structure of Posture Vector

2. Pose Similarity

Since the vector represents the human posture, the system needs a factor for comparing any two poses.

2.1 Angular Coefficient

An angular coefficient is defined as it shows the angle between two vectors. Specifically, the following was defined as the angular coefficient because the vectors' inner product includes cos theta.

$$Angular\ Coefficient = cos\,\theta = \frac{\vec{A}\cdot\vec{B}}{|\vec{A}||\vec{B}|}$$

This angular coefficient has a value between 0 and 1.

2.2 Magnitude Coefficient

However, several experiments have found that the magnitude of the angle between two vectors is negligible when comparing two vectors with similar postures and different sizes. Thus, the angular coefficient only works significantly in some circumstances.

Therefore, to make the pose similarity more sensitive, it was necessary to multiply a coefficient proportional to the vector's magnitude. A normal distribution-shaped function is considered that converges to 1 as the size of the two vectors is similar, and 0 as the size difference is more prominent. In this manner, the magnitude coefficient is defined as follows.

$$Magnitude\ Coefficient = e^{-\frac{(|\vec{A}|-|\vec{B}|)^2}{|\vec{A}||\vec{B}|}}$$

2.3 Vector (Pose) Similarity

In conclusion, the following vector similarity is defined.

$$Vector\ Similarity$$
$$= A.C \times M.C = \frac{\vec{A}\cdot\vec{B}}{|\vec{A}||\vec{B}|} \cdot e^{-\frac{(|\vec{A}|-|\vec{B}|)^2}{|\vec{A}||\vec{B}|}}$$

3. Algorithms

In order to analyze training by looking up all the frames in real-time, a considerable amount of computational resources are needed. Since accessibility requires a manageable amount of computation from popular CPUs and GPUs, the system has to detect specific postures in a specific exercise to judge.

3.1 Pose Reference

A reference pose is needed to compare the posture seen in the video. Therefore, the starting and maximum postures are defined as references for the three major exercises, squats, deadlifts, and bench presses. That is, a total of six reference poses are set.

3.2 Determining the Type of Exercise

Suppose vector similarity between the posture shown on the camera and an exercise's starting posture exceeds a particular threshold value. In that case, the system considers the exercise to have started.

3.3 Determining the Number of repeats

Suppose an exercise is in process based on the judgment of part 3.2. Then, calculate vector similarity between the maximum and current postures. Next, record that the maximum posture is reached if it exceeds the threshold value. Lastly, add one repeats when the posture reaches the start position again.

3.4 Determining the Number of Sets

If the number of repeats is greater than 0, and the time with no action is greater than the time threshold, add one set and initialize everything.

3.5 Algorithm Flowchart

Figure 5 shows the algorithm described so far.

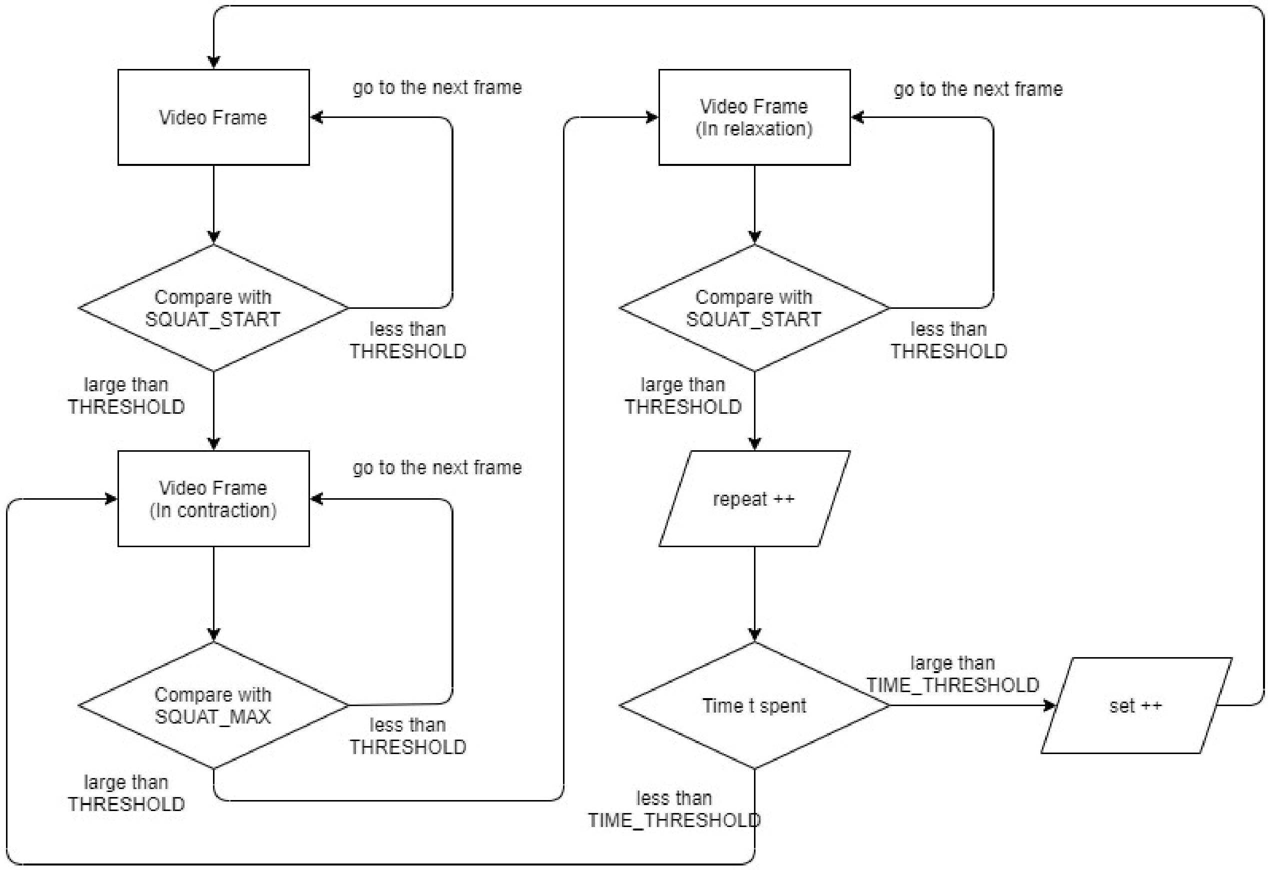

Figure 5 Algorithm for Squat

The algorithm is the same for the other two exercises, with THRESHOLD set at 0.94 and TIME_THRESHOLD set at 10 seconds.

However, for bench press, key points corresponding to the lower body were excluded when calculating vector similarity. The lower body is fixed and does not move, so the similarity is too high when it is included.

III. Results

1. Results Video

During the entire training process, the program monitors the repeats and number of sets for each exercise in real-time with a skeleton that tracks posture. In addition, these processes can be outp ut as an avi file.

Figure 6 Captures of Result Video

2. Routine Text File

In addition to the video, the results can be printed into a text file. It contains how many repeats and sets of each exercise, how many sets the total number of sets, and how many hours and minutes training took.

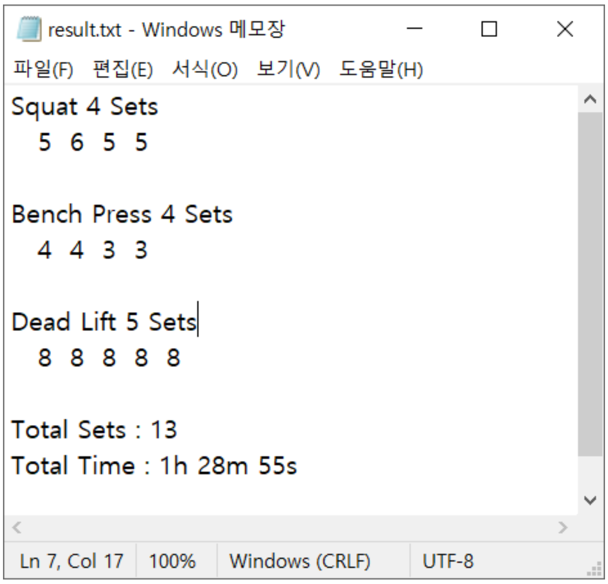

Figure 7 Result Text File

IV. Future Research and Conclusions

Using the same algorithm and program, registering only the reference pose will allow the quick addition of many other exercises.

The program will motivate weight training beginners by giving them visible results and will provide convenient routine records for middle and high-level users. In addition, accessibility is also significant because it can be applied immediately with a camera and a computer.

However, there are some problems, and the biggest problem is that it needs to detect the weight of the components of weight training. The best way is to develop and synchronize the weighing hardware on the floor, although it would be nice if computer vision could detect the weight of the barbell or dumbbell. It could be a challenge in future research.

In addition, since it was detected using only one camera, there is also a disadvantage that training is recorded with only two-dimensional coordinates. In training, detecting a person's posture three-dimensionally seems necessary to distinguish many exercises. If we continue with computer vision, there may be a way to synchronize and detect two or more cameras, or there may be a way to use other platforms such as infrared sensors or Tof cameras.

If this system is installed in a real gym, it is unrealistic for a user to carry a camera and install it on a tripod to record a routine. It is most realistic for a user to record a routine without installing a camera in each power rack or exercise device. A technology that distinguishes the user must be introduced for this concept. It is likely implemented with facial recognition and discrimination technology.

If the above definite problems are improved in future studies, it could be a better automatic exercise routine recording system.

Reference

- openPose GitHub

https://github.com/CMU-Perceptual-Computing-Lab/openpose

- Apple Watch

https://www.apple.com/kr/watch/

- Samsung Galaxy Watch

https://www.samsung.com/sec/watches/galaxy-watch3-r840/SM-R840NZKAKOO/

- Nintendo Switch

https://www.nintendo.co.kr/switch/specs/index.php

- XiaoMi MiBand

https://www.mi.com/kr/mi-smart-band-5/specs/

- OnFit Mirror

http://www.onfit.com/product/OnFitFitnessMirror/