Title	How to adjust workout posture to correct position by using electric equipment
Sub Title	
Author	肖, 俏(Xiao, Qiao) 岸, 博幸(Kishi, Hiroyuki)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2022
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2022年度メディアデザイン学 第936号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002022-0936

Master's Thesis
Academic Year 2022

How to Adjust Workout Posture to Correct Position by Using Electric Equipment

Keio University
Graduate School of Media Design

Qiao Xiao

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Qiao Xiao

Master's Thesis Advisory Committee:

Professor Hiro Kishi	(Main Research Supervisor)
Professor Akira Kaito	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Hiro Kishi	(Chair)
Professor Akira Kato	(Co-Reviewer)
Professor Matthew Waldman	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2022

How to Adjust Workout Posture to Correct Position by Using Electric Equipment

Category: Design

Summary

It is important for workout beginners to imitate the poses of experts in an absolutely correct way. This is especially important when they work out by using YouTube training videos, since the training effect depends greatly on the pose gesture. However, it is difficult for beginners to learn the proper poses for all patterns from watching videos since the detailed angle of every movement is completely different. The workout system is proposed to measure the user's 15 body joints, analyze 2D pose images, process and compare input body joints and output body joints. Our system simulates the position of the user and overlays that onto the video of the trainer. When the system detects any difference between the user and trainer's position, it will accordingly give audio or visual arrow feedback to remind the user to lift up or lower their body. The thesis analyzes if the audio, arrow, audio/arrow feedback work the best on improving the workout effect. Also, comparing the feedback that is given to the leg and arm separately, and given to arm and leg collectively. The more effective way of giving feedback is also tested. Results of the research has potential applications in creating new forms of re-programmable body movements in both virtual and physical environments.

Keywords:

body movement, gesture tracking, image identification, camera capture, human-centered computing, feedback system

Keio University Graduate School of Media Design

Qiao Xiao

Contents

List of Figures

Acknowledgements

I am indebted to Professor Hiro Kishi and Professor Akira Kato, for their kindness in guiding me and encouraging me to challenge myself during the research. I want to thank my mentor, Hiro Kishi, for meeting with me and guiding the research. Thanks for your understanding when I wanted to postpone one semester to graduate due to mental reasons, you showed me great support to help me go through the difficult time. Professor Kato also gave me the wise advice I needed and kept guiding the direction of the research. I appreciate his help with ideation for the prototype, experiment design, the slide layout and format change, your detailed guidance is the key to the success of the thesis. Also,the members of Creative Industry's laboratory gave me insights on prototyping.

Thanks for Matthew Roach for being cheerful and helpful all the time. You inspired my idea of improving the experiment effect and always there for me. I am also grateful for Meng Xiaru and He yan to instruct me how to design questionnaire and what the content I need to pay attention to.

For the coding task, I have to appreciate HuangWeihan. You helped me set the table, remove unnecessary quote, and insert special symbols. Your assistance have saved me a great deal of time to do coding.

I appreciate the participants of the experiment (Jiang Ning, Pan Jie, Cheryl Zhang, Fang Huan, Fumito Tsuji, Kakinuma Genta, and the friends you brought to do the experiment), without you, I couldn't get the user study done. I also want to thank Genta san for doing Japanese grammar check for me, so I can translate the questionnaires to Japanese. Many people helped me during my research. Your kindness gave me unfailing hope to finish this thesis.

Lastly, I must express my deepest gratitude to my parents for supporting me financially and mentally during my study. They support my decision to further my study abroad, and a lot of unexpected things happened. Without you, I might not have the courage to finish my studies in Japan.

Chapter 1
Introduction

In this chapter, the background and the brief of the research will be explained. First, the situation of workouts prior and during COVID-19 will be discussed. Secondly, the limitation of the current solution and the new improvement plan will be introduced. Third, the future application of the workout plan is also included. Last but not least, the thesis purpose, contribution, and the thesis outline will be shown.

1.1. Motivation: Evolution of Workout Method

COVID-19 has had a significant impact on numerous enterprises, particularly in the fitness sector. People who are unable to go to the gym spend months cooped up at home doing very little exercise. Some people connect to the anticipated explosion of online courses to find solutions to their concerns at home. However, some people may find it difficult to choose the best strategy for getting the best workout outcomes. A traditional workout is focused on receiving instruction from certified fitness athletes and trainers in person. How to design the best workout for at-home strategy is the obstacle that every fitness enthusiast must face.

1.1.1 Situation of Workout during COVID-19 Era

Mask use during exercise causes people to feel uneasy, which reduces the number of people who may visit the gym and increases the possibility of contracting an infection. Despite all of these drawbacks, the retention rate for gym members remained poor, as evidenced by the fact that "group exercises have a 56 percent higher likelihood of having members quit and cancel their memberships" [1].

1.1.2 Situation of Workout after COVID-19 Era

People's ideas haven't changed, even though some gyms have adjusted and adapted to new procedures to stop the infection from spreading in the future. Globally, 27.52% of gym patrons and 34.94% in the US don't plan to return after receiving the vaccination. Once their loved ones, family, and friends have received vaccinations as well, another 24.24% will return [2].

According to ClubIntel's analysis, 57% of the respondents who stopped coming to a gym gave this reason for stopping: they didn't feel COVID-19 was enough under control to justify going back [3].

People are looking for alternatives to the gym because gym-goers no longer frequent them. Due to COVID-19, 40% of respondents performed their first workout at home (Harrison Co., May 2020). Watching YouTube videos is the most common way to exercise. Virtual workouts continue to have a substantial market even as some studios open their doors, according to data that reveals users of the workout apps usage are almost 20% more per month than they were last year [4].

After using the home workout services, people find out a lot of advantages on work out at home, such as lower cost, time-saving, not being affected by bad weather, etc. The statistics also supported the phenomenon: it shows that 60% of gym patrons appreciated their home exercises so much that they intend to resign from their memberships [5].

1.1.3 Limitation of the New Solution

Unlike sessions at the gym, where consumers might receive expert advice from a trainer or their workout partner. There is no sufficient instruction provided by the teacher when people exercise at home. Lack of exercise knowledge might result in errors that harm the body.

Google searches for knee pain have jumped by 471% in the month and a half since Monday, March 23, while those for sprained ankles and low back pain have also increased by 267% and 157%, respectively [6]. Because many people practice workouts at home, such as lunges and squats, incorrectly. Because of people's bad technique, which puts additional strain on their joints, online inquiries for press-up pain reached a 12-month high [7]. As the data showed, lack of professional guidance in-home workouts has caused different kinds of physical pain [8].

1.2. Overview and Goals

In this thesis, we present Home Workout, a Python program, which offers correction suggestions to realize better workout effects. By comparing the positions of users to trainers, users will be given corrective feedback if their position is wrong, and accordingly, move their positions to the correct place. For the feedback given to participants, audio and arrow feedback are the two main feedback given to participants. Also, the feedback is given in two different ways: separately to leg and arm, or given collective to leg and arm. There are two main goals to achieve:

The first goal is to compare if the feedback can improve workout effect, if yes, then test to compare the audio, arrow, arrow and audio feedback to see which one is the most effective and which one is the least effective.

The second goal is to compare if giving feedback separately to leg and arms and giving feedback collectively to leg and arms will cause any difference on improving workout effect. If yes, then test to see which workout method is more effective.

The home workout program can recognize the user's body movement by analyzing 2D pose images, processing, separating, and comparing the user's 15 body joints and trainer's body joints, and then providing guidance to correct the difference.

The data of user's body movement can be uploaded to the python system, and the user's data and trainer's data will be compared. When the difference of body joints is greater than 10 degrees, the device is going to judge the position as the wrong position, and accordingly give different feedback to different body joints, which helps users correct their body position.

For example, when the user lifts her leg slightly lower than the trainer, our system will recognize the mistake by analyzing the position data from the knee and ankle, then assess it as a "Fail" and give feedback. According to the feedback suggestions to leg and arm, the user will think about where she did wrong, try to lift her leg, and reach the correct position as the trainer does.

To evaluate this system, there were three experiments related to testing and comparing the workout effects. There were 16 participants who did the experiments and filled out questionnaires afterwards. Also, we set up a camera to test the body movement of participants to see if they get improvement in the process of workout with feedback received and how many PASS (if the position of user and position of trainer is the same, then they get PASS in the system) do they get. In this thesis, the planning and designing of the experiment will be introduced.

After the experiments, the participants are free to talk about their feedback about the experiment, and we will look into how we can augment the workout experience in the future.

1.3. Contribution

The Human-Computer Interaction (HCI) perspective was merged into the Home Workout Python Program, and it can be used as a set of standards and recommendations for this study as well as future ones.

A revolutionary program that acts as a display and provides users with visual cues is introduced in the program. The Home Workout Program's design, execution, and applicability in real-world situations are all covered in this thesis. One software produces the desktop interface, camera recognition, and comparison analysis together.

The following is a list of the thesis's contributions:

I put forth a prototype that can analyze a 2D image using a variety of methods and tools. It also features a model that enables real-time movement to be continually recorded, and the resulting body pose data can be compared. Additionally, the error range of the angle difference between the participant and the trainer can be altered.

According to our evaluation (N=16), I discovered that delivering audio and arrow input simultaneously while also providing feedback for both the arm and the leg delivers the best training effect.

I discussed the design principles and their possible uses in a variety of fields, including games in addition to training videos.

1.4. Thesis Outline

This paper consists of 5 chapters, the thesis outline is shown as follows:

Chapter 1 introduce the motivation, then followed by a brief introduction for the experiment design and goal, and pinpointing the research contributions, as mentioned above.

Chapter 2 will discuss the related work, which focuses on the real-time feedback, sensors function, and platform building. Also, how the Home Workout Program differs from the other workout programs. At the end of chapter 2, thesis contributions will be stated.

Chapter 3, the ideation will be introduced. Then the concept design and layout of the prototypes will be shown.

Chapter 4 demonstrates the experiment design on two methods. The result and discussion will be revealed later on. The evaluation is another main part, which adopted various methods to analyze the data and information collected.

At the beginning of Chapter 5, the conclusion will be given. The limitation and the future improvement of the project will be addressed.

All the study materials including questionnaires are in the appendix.

Chapter 2
Literature Review

This chapter will cover the literature review and related works.

Related work will be categorized into three sections. First of all, the gesture tracking technologies are discussed. Specifically, recognizing the body joints, tracking in real-time basis, and camera setting is discussed. The second discusses what code and algebraic method is used to configure the system. Third, previous applications through camera detection and image analysis are shown.

2.1. Body Recognition

2.1.1 Open Pose System

Because human pose data is frequently needed in the health and sports field, many studies have been conducted to recognize the data. To recognize the data of a person from an image, we gain insights from body recognition techniques[1] called OpenPose [9], the first real-time multi-person system to jointly detect human body, hand, facial, and foot key points (in total 135 key points) on single images. The ability for machines to recognize people in photos and videos is made possible in large part by real-time multi-person 2D posture estimation. The researchers offer a real-time method for identifying numerous people's 2D poses in an image. The proposed method learns to link body parts with persons in the image using a non-parametric representation that we call Part Affinity Fields (PAFs). With the use of this tool, we can track a single person's gestures and visually smooth their movements.

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

Overall, the methods that are now in use can be divided into three groups: absolute pose estimation, relative pose estimation, and appropriate pose estimation, which combines the two.

The first is the absolute posture estimate approach, which is based on heat map matching, active and passive landmarks, navigation beacons, and satellite-based navigation data. The second technique is based on dead reckoning and updates the human pose incrementally by estimating the distance from a known joint, or the beginning position and orientation of a human.

Fundamentally, most algorithms anticipate a person's position in relation to the background using their pose and orientation. It is a two-step approach that first recognizes human bounding boxes before assessing the pose of each box.

Next, the major joints for a person, including the elbow, knees, wrists, and so forth, are then estimated. Depending on the purpose, we can estimate poses for a single person or a group of individuals.

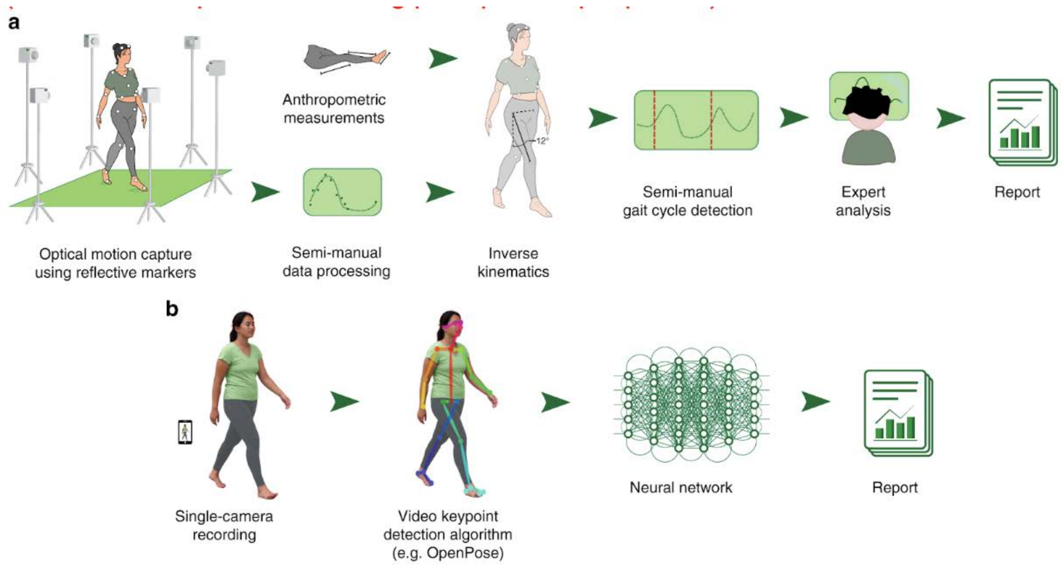
The model guesses the poses for a single individual in a certain scene when performing single pose estimation. In contrast, the model calculates the postures for several people in the given input sequence when using multi-pose estimation. (Figure 2.1)

2.2. Real-Time Gesture Tracking

A technique put out by "ExemPoser" [11]forecasts the pose of a rock climbing expert based on the climber's hands, feet, and the placement of the grips they are using. Beginners can easily use their system in real-time from smartphones thanks to the way it was developed. (Figure 2.2)

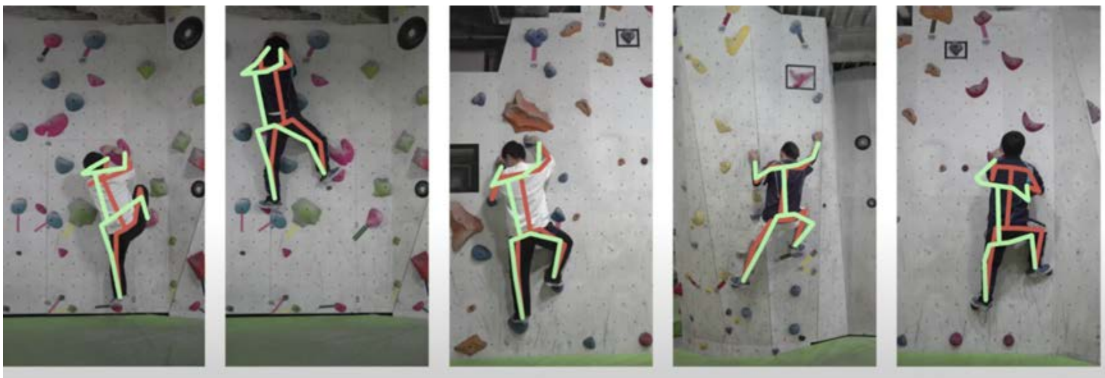
The ExemPoser has the capability to recognize real-time performance in order to realize the system. The neural network that powers the first feature has input layers with 8 dimensions and output layers with 20 dimensions. using a straightforward neural network that is fully connected. This network generally has low latency and is lightweight [9].

The prediction shows the result of the beginner's data. The red line indicates the poses of climbers, and the green line indicates the predicted poses.



(Source: https://www.researchgate.net/publication/343632899) [10]

Figure 2.1 Open Oose Working Mechnism



(Source: https://dl.acm.org/doi/10.1145/3384657.3384788) [12]

Figure 2.2 Gesture Tracking by Open Pose

2.3. Camera Setting

2.3.1 RGB Camera

Deep learning technology has recently been successfully used into at-home workouts. A camera is often used in interaction systems for at-home workouts to recognize a user's stance and action. A camera may be used by some interaction systems to detect a user's pulse. In order to utilize open-source programs related to autonomous navigation, the mobile robot is implemented in the ROS robot operating system. The system uses the RGB camera to locate anatomical key points in 2D space in order to estimate human position for a user performing at-home exercise.

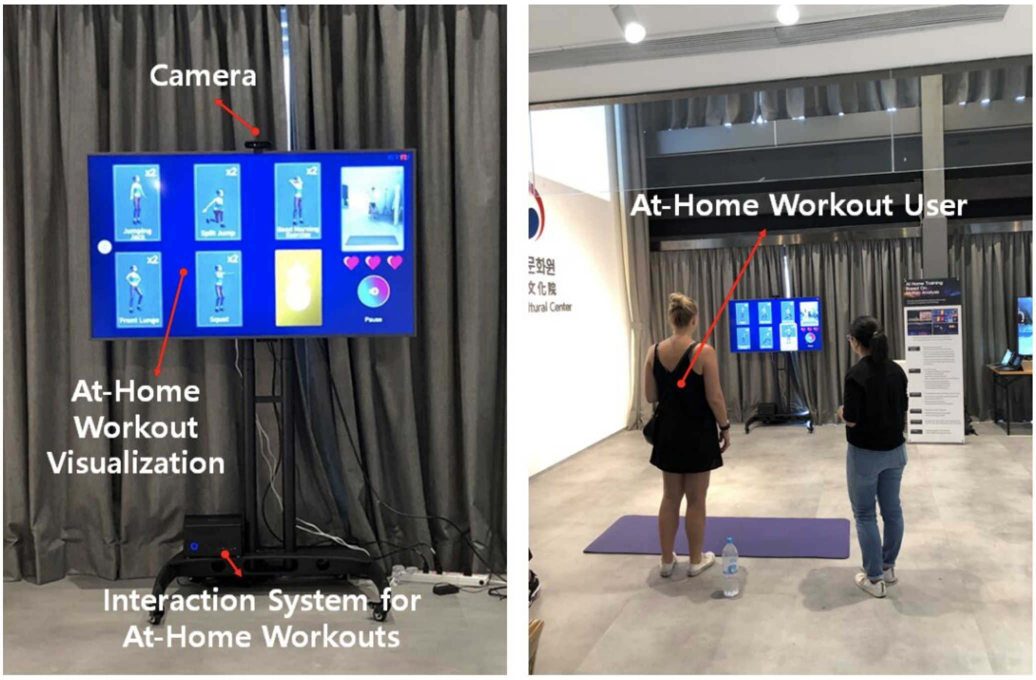
In the Mobile Robot at Interaction Systems made by Samyeul Noh [13], using the RGB camera, the system uses action recognition to identify and distinguish different activities included in video clips (a sequence of image frames) throughout the course of at-home workouts. (Figure 2.3)

Using the RGB camera, the system recognizes bio-signals by detecting the small color changes caused by pulses on the surface of human skin, particularly on the face. Below is a description of the camera's space setting [14]:

To be more precise, the unity contents visualization module offers the user a variety of graphical data for clever interaction. During a workout at home, the user's various activities are recognized by the action recognition module. The user's 2D localisation of anatomical important points is estimated by the human posture estimation module. An RGB camera and an infrared rays camera are both utilized by the bio-signal recognition module. To enhance the effectiveness of action recognition and human posture estimation in the pre-processing stage, the human data creation module generates additional training data. [13].

2.3.2 Distance Setting

As for the camera distance setting in "ExemPoser" [9], users don't need to worry about the position of the camera because the technology allows for any arbitrary distance between the wall and camera. Pre-processing makes this functionality possible. If these body joints are shrunk at the predetermined ratio, the real



(Source: https://ieeexplore.ieee.org/document/8939887) [14]

Figure 2.3 Interaction System for At-Home Workouts distance between the holds could not be taken into account because the distance between the camera and the wall of video is unknown. Next, we note that on a true scale, the distance between the neck and the midpoint of the hip is practically same. As a result, the image of the climbing attitude will take on a given size by reducing the joints at the rate of average distance [11].

2.4. Comparation Design

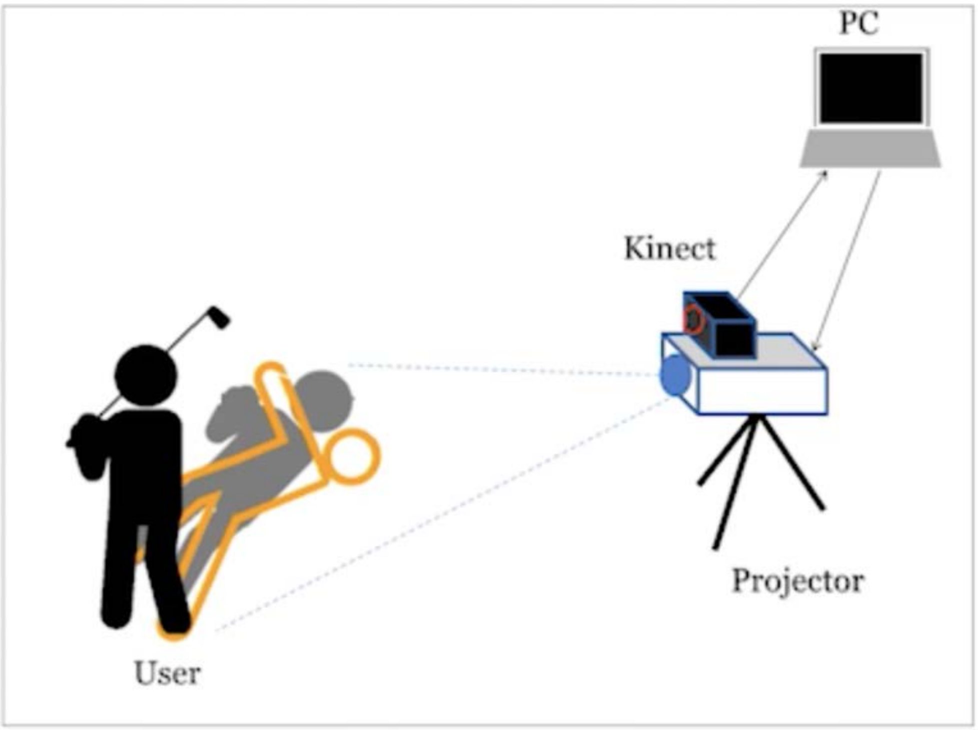
The article [15] suggested a real-time visual feedback-based golf teaching system. The device casts the user's virtual shadow on the ground in front of them, giving them feedback without causing their form to collapse. In order to make the user aware of the differences between their form and that of the expert, the contour of the expert is also superimposed over the user's virtual shadow. Moreover, the expense of adapting to this training approach can be decreased by mimicking the shadows already utilized in conventional golf instruction. (Figure 2.4) This work taught us to compare the trainer's virtual movement to the user's movement, and to overlay the movements to make the user aware of the differences.

2.5. Sports in Human-Computer Interaction

2.5.1 System Configuration

One of the main problems in engaging in exercises at home is that there is no proper guidance and feedback provided to align the exercises to the correct movements due to the absence of a physical trainer. Fitness Mate [13] discusses how a system is designed and implemented to enable users to engage in physical exercises without the presence of a physical trainer. Visual C++, Matlab programming language, and Unity game engine are the technologies used to develop the system. The main goal is to provide a home-based environment where people can engage in physical exercise without the presence of a physical trainer and to avoid adverse physical injuries.

The "ExerCube" [16] is a fitness game setting for adults, which affords immersive game play experiences while engaging in a playful motor-cognitive and



(Source:https://dl.acm.org/doi/abs/10.1145/3279778.3279927) [15]

Figure 2.4 System Configuration of Golf Training coordinated functional workout. Some exergames are developed from a mere entertainment trend to serious training applications, such as Nintendo Wii and Sony Move. Innovative technologies are introduced to turn the living rooms into playful training settings by offering virtual training simulations. The combination of training concepts and game design helps people to keep fit in a motivating and attractive way by putting the user experience at the forefront of the design and evaluation process.

To address the limitation of the camera's mobility issue, researchers presented the system to assist people with working out at home through various advanced deep learning technologies [17] such as action recognition, human pose estimation, and bio-signal recognition. We learned from the modules used below: unity contents visualization, action recognition, human pose estimation, bio-signal recognition, and human data generation.

2.6. Precedent Design on Digital Fitness Platform

In the following section, two types of feedback were introduced. One is giving feedback by trainer. The other one is automated feedback by the app itself. These apps/tools show the possibilities of giving different types of feedback. The section below will cover their distinct functional features, out-standings, and short-backs.

2.6.1 OpenFit

Feature: This App helps to expand the platform's offerings to its subscribers, including the addition of live classes, where certified trainers provide real-time feedback to participants through the optional use of a phone's camera, along with personalized meal-prep and nutrition tracking tools.

Pro: Users do not only have access to live classes — they can also stream results-driven, on-demand workouts led by world-class trainers. This app also offers Live interactive classes - user can schedule the classes.

Con: No arrow feedback given to users, and audio feedback is not precise enough as it is not given by AI tool. The automated feedback can also be offered at anytime without the hassle of scheduling classes.

2.6.2 All-in-One Machine (Tonal)

TONΛL

Feature: Tonal lets you take a variety of classes such as dance cardio, HIIT, Pilates, and yoga. Tonal analyzes the user's form and gives feedback when it detects opportunities for improvement.

Pro: Offer challenging workouts in the privacy of your home. Its resistance cable system that allows users to easily learn to strength train at home and lift up to 200 pounds

Con: User needs to purchase the hardware and install it in your home or apartment. This is not convenient to have everywhere.

2.6.3 Certificated Trainer (Mirror)

MIRROR

Feature: Exercising in front of a fitness mirror helps you check and adjust your form.

Pro: Mirror's Bluetooth capability also extends to heart-rate monitors, Apple Watches, and other smart technology. This can provide up-to-date feedback on your physical condition to the Mirror display. Also, The feedback of calories burned and heartbeats per minute indicated on the screen.

Con: There is no arrow or audio feedback to help change the position. The only way for user to make improvement is observing their positions in the mirror.

2.6.4 Performance Tracker (Keep)

Feature: Keep App has detailed performance tracking of history record to provide customized report.

Pro: Keep offers beginner-friendly data report to track how many hours they workout and what classes they use. User can also hear audio feedback to remind them what is to keep in mind while workout.

Con: There is no visual feedback as arrow, and the audio feedback is not customized to the real time workout.

Chapter 3
Concept Design

Workout people need the feedback to improve their body gesture and realize better workout effects. In addition, COVID-19 has forced people to do workout at home, which lead to the difficulty of receiving feedback from trainers in the gym.

In order to overcome the difficulty of not being able to receive feedback and not receiving the optimal feedback, this research researched different methods of giving feedback, and the way of giving feedback to multiple or singular body parts. The technology get evolved to observe, receive, and give is explained in this chapter.

3.1. Problem Definition and Our Mission

COVID-19 has forced people to stay home, thus they can only do workouts at home instead of going to the gym. Since most people are not professional trainers, what they can do at home is imitate movements from YouTube videos. To know the poses of the experts, beginners need to look at the example of experts for each scene. Besides COVID, there are also other reasons leading to the gym quitting behavior, which include the high pricing, not fitting in the environment, and moving away.

Also, the elderly population in Japan has been rapidly increasing, so seeking a cheap option for maintaining their health condition has become an essential issue. Workout at home has become the choice the elderly will go for. In the big data era, these common barriers are easily addressed by offering digital platforms. By working out at home, people don't have to worry about the price, social distancing, not fitting in with others, and problems like that.

3.1.1 Observations and Interviews with Gym Goers

This section will talk about how the concept design is developed from the initial idea of augmenting workout effects. Initially, through the field work of going to gym which is next to where I live about 3-4 times a week. To make the workout experience more fun, I used several YouTube trainer's video to do workout. The most frequently used one is Pamela's workout video. The reason why I favored to use Pamela's video (Figure 3.1) is because it doesn't give extra audio instruction, but gives clear visual images. While not knowing if my position reached to the perfect place, I feel more concentrate to Pamela's body movement by merely looking at it. Therefore, I raised the question that if only giving visual feedback will help people realize better workout effect, or give audio feedback will do the job.

As for the visual feedback, I thought about giving different forms of feedback, such as arrow feedback, line feedback and etc. After asking for five friends' opinion about which form of feedback I should give, the line with arrow attached at the end receive the most recognition on giving the best workout effect.

Also, every time when I adjust my body positions, I would start from correcting one body part, then go correct the next body parts instead of adjusting several body parts together. This can be also related to the method of giving feedback. Will it be more effective to give feedback on leg and arm separately, or will it be more effective to give feedback to both leg and arm collectively.

3.1.2 Fieldwork at Gym

To further prove the design can solve people's pain points. We conducted several pre-interviews to my friends who go to gym often. The interviewee is asked to answer about their workout experience while using workout video. First case is to do workout with audio feedback given, second case is to do workout with themselves being seen in the mirror to compare, which is the visual feedback to learn from. After the workout, feedback from the experiment subjects will be collected to evaluate if the experiment topic is meaningful and helpful on improving the workout positions. The feedback is through asking about the feeling and rate the workout experience between the two workouts, which is based on

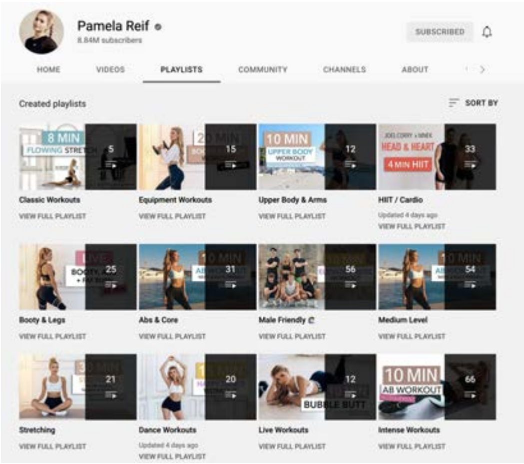


Figure 3.1 No Audio Feedback Given by Pamela

qualitative review. Also, they were asked if they focus more on correcting one body part first and then move to next, or focus on correcting several body parts together. Which body parts to focus first in the workout process is also asked. Through comparing the workout effects between the two feedback and normal workout without any feedback, and comparing feedback give to body parts in sequence or all together, we reach the conclusion that giving feedback is necessary and how to give feedback should be the key point to discuss. (Figure 3.2) (Figure 3.3)

Figure 3.2 Fieldwork at Gym

Figure 3.3 Fieldwork Setting at Gym

3.2. Goal Setting

Understanding what is the best way to give feedback during workouts. By doing individual position movements (doing one position at a time, such as: doing leg first, then do the arms) or group positions (doing multiple positions at a time, such as: doing leg and arm at the same time) and receiving different feedback (visual feedback, audio feedback, audio and visual feedback together). What is the best feedback and how much feedback given at one time can help workout users to realize the best workout effect.

3.3. Prototype Design

3.3.1 Core Mechanism

First, the image of the workout user is captured by a fixed point camera from behind. At the moment the user is stationary, the user's body joints are captured and calculated from the image using an open pose. The positions of hands and feet are extracted from body joints and our device receives these positions. The system can compare the poses of the user and trainer and give suggested advice that can further promote beginners' skill improvements.

Package of the code includes PyQt5mediapipeopencv-pythonnumpy.

To create a GUI-based program, PyQt is chosen to enable widgets being built to create complex GUIs. In Qt (and most User Interfaces) "widget" is the name given to a component of the UI that the user can interact with. User interfaces are made up of multiple widgets, arranged within the window. [1]

MediaPipe, as a Framework for building machine learning pipelines for processing time-series data like video, audio, etc. For the human pose estimation, MediaPipe can form the basis for yoga, dance, and fitness applications. It can also enable the overlay of digital content and information on top of the physical world in augmented reality. MediaPipe Pose is a ML solution for high-fidelity body pose tracking, inferring 33 3D landmarks and background segmentation mask on the whole body from RGB video frames utilizing our BlazePose research that also powers the ML Kit Pose Detection API. [2]

To add support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays, NumPy has been added to the package.

OpenCV is used to solve computer vision problems. Computer vision includes understanding and analyzing digital images and processing the images or providing relevant data. (Figure 3.4)

RBG camera is also used for creating orthomosaic maps that show your entire field at once, and they can capture aerial videos.

3.3.2 Visual Presentation

The MacBook's built-in camera is used to collect the video and import it to the computer system. The GUI interface present the interactive visual components, which include the buttons "Upload Video", "Upload Action Picture", "Start", "Pause". The trainer's video is shown as "Standard Action (Trainer)". (Figure 3.5)

---

1 `https://www.pythonguis.com/tutorials/pyqt-basic-widgets/`

2 `https://google.github.io/mediapipe/solutions/pose.html`

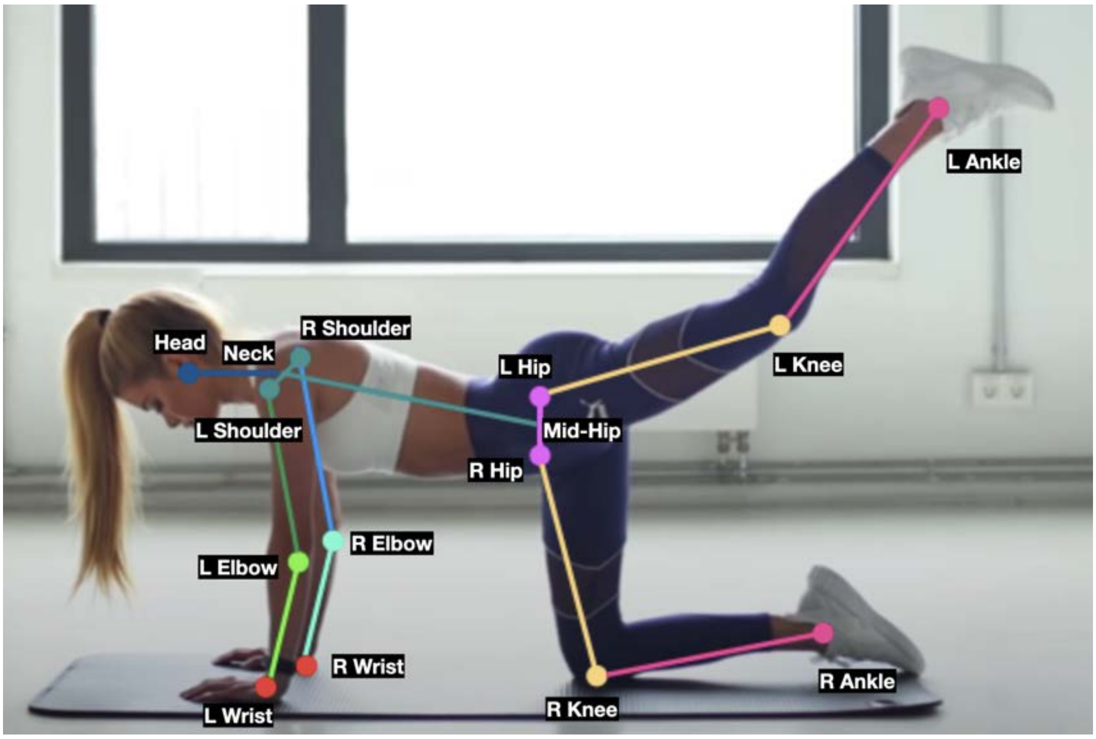


Figure 3.4 Body Joints Recoginition by Open Pose

Figure 3.5 Visual Presentation of GUI Interface

3.3.3 Gesture Recognition

The device compares the input and output data of the two-dimensional (2D) position of body joints extracted from the image of the workout users. To detect human pose, we used OpenPose to detect the human body. Open Pose estimation is a computer vision task that represents the orientation of a person in a graphical format. In the first step, the image is passed through the baseline CNN network to extract the feature maps of the input In the paper. In this paper, we use the first 10 layers of the VGG-19 network. This real-time multi-person system is used to jointly detect human body, hand, facial, and foot key points [18]. Then, when an image or video is given to the pose estimator model as input, it identifies the coordinates of these detected body parts and joints as output and a confidence score showing the precision of the estimations. Despite that OpenPose can extract 135 body joints, here we only cover 18 body joints. The rest of the body joints are not considered, as it has little impact on the workout effect.

3.3.4 Comparison Design

As everyone has a different lengths of limbs, instead of measuring the length of limbs, we measure the angle of the body joints.

The positions of the user's body joints correspond to the trainer's body joints. The input user's body joints will enter the system and be used to compare with the trainer's body joints. Accordingly, if the body joints don't fit, the system will indicate ERROR on the screen, otherwise, the screen will indicate OK.

The "threshold" function is used to define when to set ERROR and when to set OK. The smaller the threshold number is, the more precise the system will be. We set the standard threshold number as 10 in the system, which means the allowable analytical error is 10 degrees. In this paper, if the difference in body angle between trainer and user exceeds 10, the system will give an "ERROR" order, otherwise, the system will give a "OK" order.

For example, if the user's left leg is 30 degrees to the main part of the body, but the trainer is 60 degrees. Our device will judge the 30-degree difference and give "Error" feedback.

3.3.5 Feedback System

The feedback would be the visual effect and sound feedback. The visual effect is to show the user whether or not they successfully performed the movement. The visual feedback would only appear if the user did the movement correctly in the next run. The audio feedback is to correct the user by telling them how to change their position to reach the correct angle. (Figure 3.6)

Figure 3.6 Illustration of Arrow Feedback

3.3.6 Prototype Working Mechanism

The overall explanation of programming process is shown below: (Figure 3.7) Be noted The details of the code is written in Appendix. After running the python code in Pycharm software, it will indicate an UI interface. The User can upload the picture to the interface, and click "Start" to compare user's position to trainer's position. When the position is the same, the screen will show OK, otherwise, it will show ERROR as we mentioned in the comparison design chapter.

To make this system work, we used code def getPts to enable the picture to be uploaded into the system, the trainer's picture size is limited to be the setting size, so that the system can compare the picture and video. The way to set up QMain windown and QDialog is through several functions:def retranslateUi, def setup Ui, def connectlnit. After the file path is set up, the user can put the video and picture into system. Here the RGB Camera, Openpose and PyQt contrainer is used to detect image. To compare the trainer's image and the workout video, the flag function is used to judge if the picture is the same or different. The difference of these two is allowed to be within a range of value, which can be changed in the code.

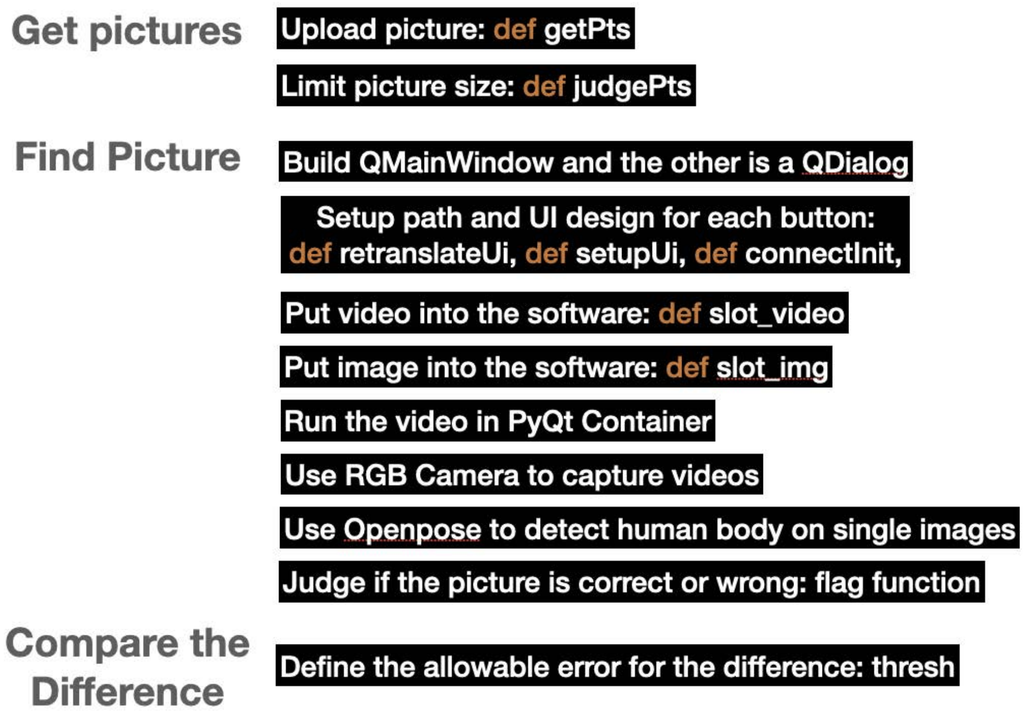


Figure 3.7 Programming Process

Chapter 4
Proof of Concept

In this chapter, firstly we give an abstract of experiment content, then the environment setting and experiment requirement is introduced. Afterwards, the evaluation method we used on measuring the experiments is discussed, followed by the according result.

4.1. Abstract of Experiment Procedure

There are 16 participants (female=8; male=8) aged from 22 to 29 years old (average age = 26.7) joined in a 60 to 90 minutes user test. The participants were asked to read through the consent form and photography permission form before the experiment starts. The experiment will start only if the participants agree with the experiment contents and requirements and sign their signature and print name. Afterwards, we explain the overall instruction of the goal and process of experiments, and show the positions they need to perform. The participants are asked to choose one set of positions, while making sure each set of positions have at least 5 participants.

After the first experiment, we will give instruction for the second experiment. Also after second, third and fourth experiments, we will give instruction and send questionnaires for participants to fill out.

Any machine sound made and questions asked in the process will be recorded.

After all the experiments are finished, we will conduct interviews with participants to see if they have any advice or confusion regarding the experiments.

Below is the picture of the participants while doing experiment:(Figure 4.1) (Figure 4.2)

Figure 4.1 Participants in Experiment

Figure 4.2 Giving the Instruction to Participants

4.2. Environmental Setting

The diagram of the environmental setting (side view) is shown below:

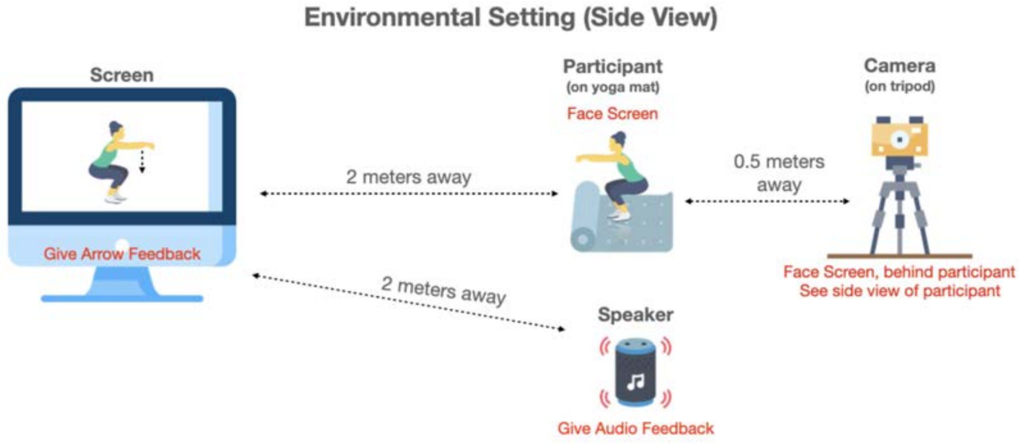


Figure 4.3 The Diagram of Environment Setting (Side View)

The side view and aerial view of the Actual Environment Setting is shown below: (Figure 4.5) (Figure 4.4)

4.2.1 TV Setting

The TV screen is a Sony KJ-49X700, with 4K 3840 x 2160 LED Panel, HDR10-Ready via Future Firmware Upgrade, using Motionflow XR 240 Technology and Screen Mirroring Technology. This 4K TV has a 48.5" LED display and incorporates Sony's Motionflow XR 240 technology.

The space between the TV screen and the participant is 2 meters, so the participant can clearly see the whole body of the trainer's workout video. To present the yoga picture and instruction guideline on TV, we connect computer to the TV using an HDMI cable.

4.2.2 Speaker Setting

To give audio sound, the SoundLink Mini II Special Edition is adopted to deliver full, natural sound. The Dimensions of the speaker is 2.1" H x 7.1" W x 2.3" D.

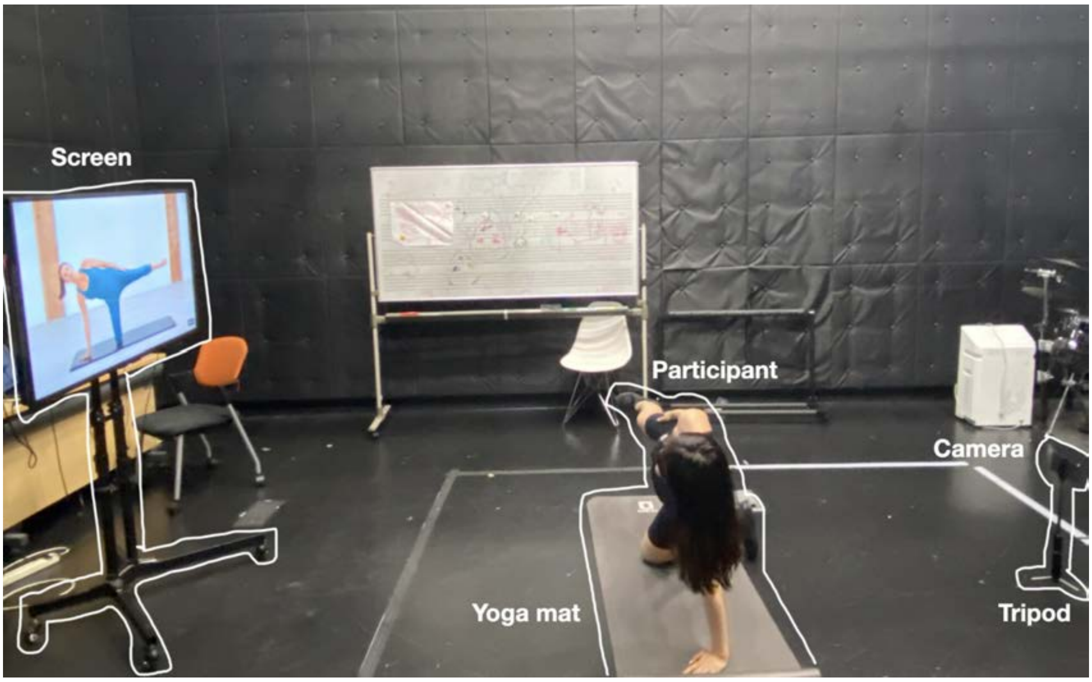


Figure 4.4 The Actual Environment Setting (Side View)

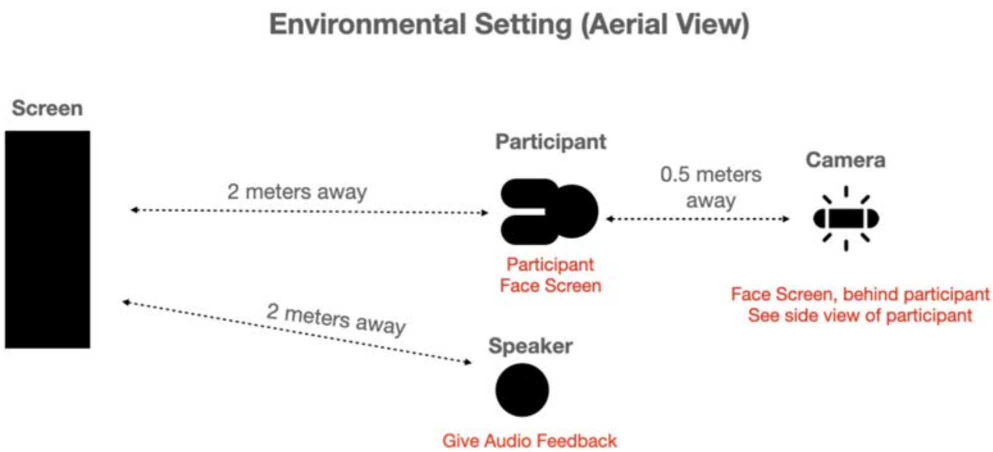


Figure 4.5 The Actual Environment Setting (Aerial View)

In this experiment, we provide one speaker and the space between the TV and speaker is 2 meters, leaving the speaker 0.5 meters away from the participant. This distance gives a good audio effect on playing the feedback sound.

Also, to make the yoga experience comfortable and fun, the Bluetooth speaker is also used to play yoga music.

4.2.3 Camera Setting

iPhone 11 camera is used to record the video and images of the users while the experiment is ongoing. (Be noted the whole experiment, including the break time and instruction time will be recorded, as there might be confusion or feedback when the experiment is in progress. The audio will be listened to later for future reference for improvements).

It includes an f/1.8 6-element 12-megapixel wide-angle lens (26mm focal length) and an f/2.4 5-element 12-megapixel ultra-wide-angle lens (13mm focal length),. The camera is set to face the TV screen, so it can take photos and videos of the user.

To stabilize the camera, we put camera on a tripod. Here the tripod we use is Magic Realm OTH-AB201, the product weight is 209g, collapse size is approximately 198mm, expand size is approximately 740mm, product size is 32X32X198mm, connection is through Bluetooth.

The camera will be used to take video for each session, so the angle of the body parts of the participants and trainer can be manually compared. The accumulative changing of the body and the ultimate position they reach will show if the feedback helps improve the workout effect and when the improvement happens.

As the trainer's positions are all taken from a side angle, to compare the image of participants to the user, the camera is also set to face to the side view of participants. The camera is set behind user in order to record the body position of user and the trainer's position on the TV for the later confirmation for which video is played for workout training. Be noted, that the camera can't be placed in the front as it can't record video of the trainer, and you will lose the function of the later reconfirmation. Also the camera can't be mounted the display as the position is faced to the screen, which can't be observed well from the angle above. (Figure 4.6) (Figure 4.7)

The participant's work out area must be directly between the TV screen and the camera, the participant's body should be parallel to the screen, so he/she can more easily imitate what the trainer is doing.

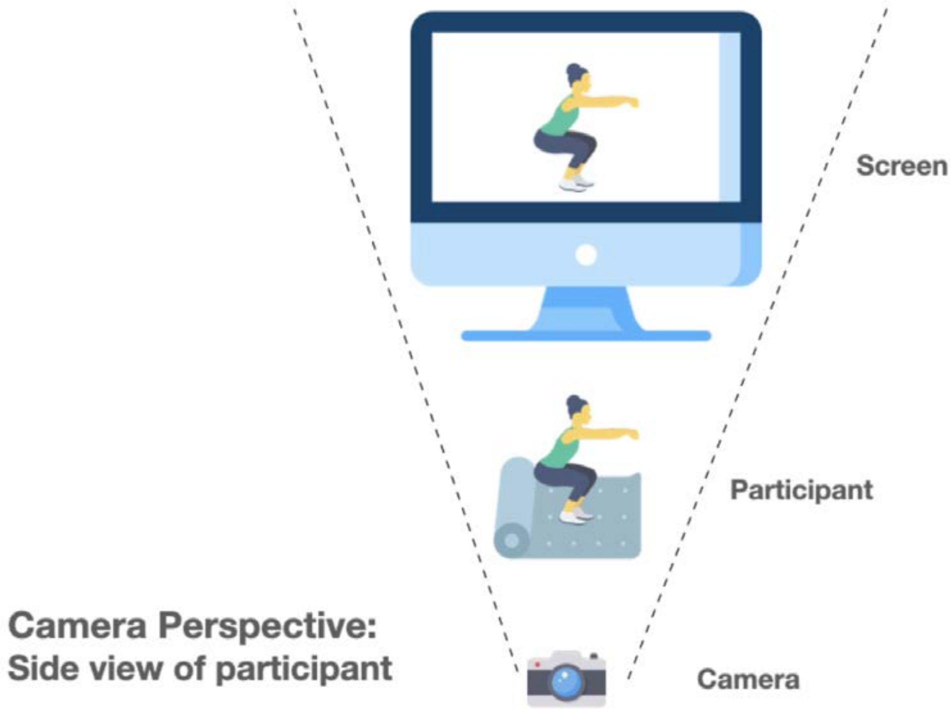


Figure 4.6 The Diagram of the Camera's Perspective

4.3. Experiment Requirement

As the software is to help improve the yoga positions, the professional might not feel their positions get significantly improved as they already can reach the perfect positions. Therefore, the target group is yoga beginners who can't imitate the yoga trainers' position correctly by themselves, thus feedback is needed in this setting. All the participants have little workout and yoga experiences.

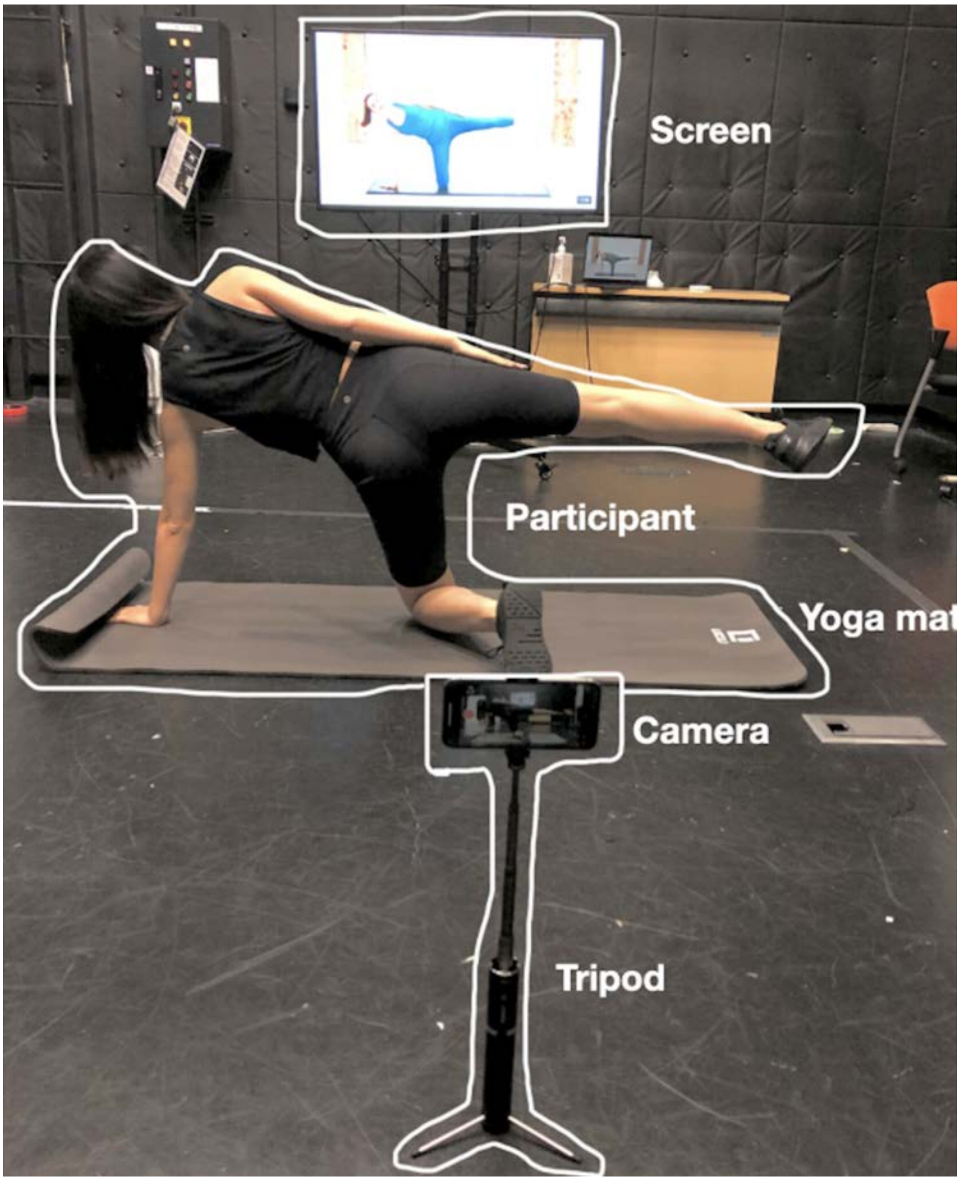


Figure 4.7 The Actual View of the Camera's Perspective

4.3.1 Preparation Work

The yoga mat is provided, as some positions are done on the floor. Yoga mat is the place for participant to stand on. Besides the yoga mat, the participants are asked to wear sportswear, in case some participants don't wear the sportswear, we prepare yoga dress for them to wear.

For the time reminder, here we use the "Repeat Alarm" App to ring the bell to tell the user the session is over.

Prior to starting the experiment, the participant will review and sign the "Explanation and Consent", and "About Photography" agreement. This has to be prepared in advance to the experiment.

"Explanation and Consent" are to get the approval to the handling of personal information and data, give participants free rights of participating, and give Intellectual property rights to researchers in Keio University or collaborators outside Keio Gijuku. Also to give the declaration of the side effects.

"About Photography" agreement is to declare the usage of the picture and videos were taken before, during, and after the study.

Be noted the agreement and survey needs to be printed out before experiments for participants.

To make sure participants are on the same page about what they need to do for the positions, the manuscript of instruction for the whole set of positions and each position is written. This way there won't be any misunderstanding between individuals to cause bias.

After signing the agreement, the participants are required to do 1 set of positions, which consists of 3 positions.

4.3.2 Positions Required

The diagram overview of the whole 3 sets of positions: (Figure 4.8)

Because not all workout positions have the same level of difficulty for each participant, we provide them with 3 options to decrease the bias caused by the different difficulty level.

Also, as all the participants have different levels of body flexibility (some participants might find it's hard to conduct a certain yoga position), we give different

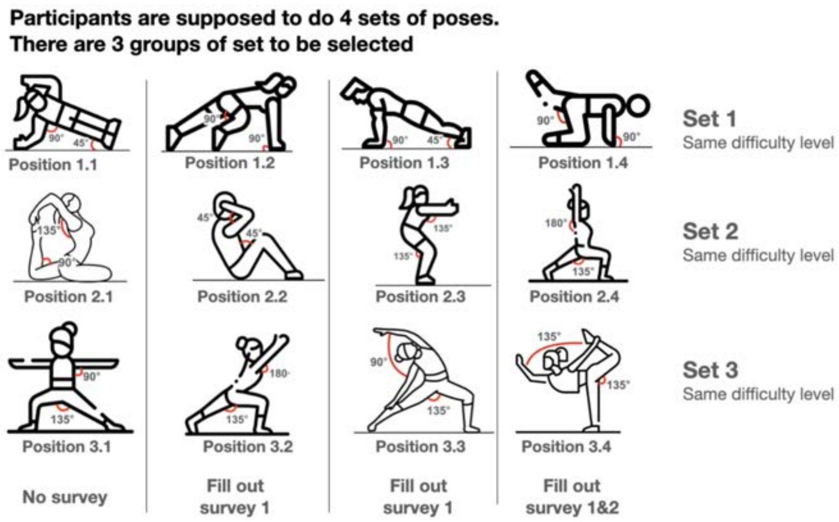


Figure 4.8 The Diagram Overview of the Whole 3 Sets of Positions

sets of positions for participants to choose.

Based on the participants' preferences, the participants are equally distributed to 3 groups, and asked to do workout positions that are designated for each group.

Everyone's set of positions include 4 different positions. Each set of positions has multiple angles to measure, although some angles are more crucial to achieving the workout positions than the others, for the convenience to compare angles between different positions, we only choose the two most important angles to measure.

To evaluate which feedback works the best, instead of dynamic positions, we will use static yoga positions to let the user have better feedback. The muscle burn is easily felt so that the participant would know if they realize a better workout effect.

The video of participants' body movement will be recorded, and uploaded to the python system we designed. The participants' s body movement will be compared to the trainer's standard image. If the difference of degree between these two is less than 10 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign. By calculating how many positions they get a OK, we can make judgment on how effective the workout is.

Below are the explanation of the 12 workout positions and the evaluation criteria for them: (Be noted every set of poses should go in a different sequence, to decrease the effect on muscle memory.)

Position1.1: (Figure 4.9)

The left arm is 90 degrees to the body, and the leg is 45 degrees to the floor.

If the degree of The left arm is 80 to 100 degrees to the body, and the leg is 35 to 55 degrees to the floor, the system will give the OK sign, otherwise, it will give the ERROR sign. (Tolerance for the angle is 10 degrees)

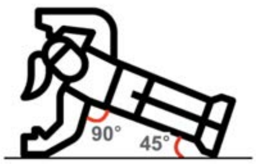


Figure 4.9 Position1.1

Position1.2:(Figure 4.10)

The left leg is lifted up, the thigh and lower leg are 90 degrees, and the arm is 90 degrees vertical to the floor.

If both thigh and arm is between 80 100 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

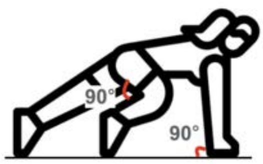


Figure 4.10 Position1.2

Position1.3:(Figure 4.11)

The third position is a push-up yoga position. The arm is vertical, and the leg is 45 degrees to the floor.

If the degree of the arm is between 80 and 100 degrees, or the degree of the leg is between 35 and 55 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

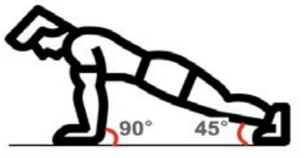


Figure 4.11 Position1.3

Position1.4:(Figure 4.12)

The left leg is 90 degrees to the right leg, and the arm is 90 degrees to the floor.

If the degree of the right leg to the left leg and the degree of the arm to the floor is between 80 and 100 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

Figure 4.12 Position1.4

Position2.1:(Figure 4.13)

The upper left leg and lower left leg is 90 degrees, and the arm and upper body is 135 degrees.

If the degree of the upper left leg and lower left leg are between 80 and 100 degrees, and the arm and upper body is between 125 and 145 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

Figure 4.13 Position2.1

Position2.2:(Figure 4.14)

The leg is 45 degrees to the upper body, and the upper arm is 45 degrees to the lower arm.

If the degree of the leg to the upper body and the degree of the upper arm to the lower arm is between 35 and 55 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

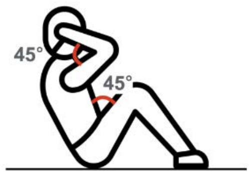


Figure 4.14 Position2.2

Position2.3:(Figure 4.15)

The arm is 135 degrees to the body, and the upper leg and lower leg are 135 degrees.

If the degree of the arm, and degree of the upper leg and lower leg are between 125 and 145 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

Figure 4.15 Position2.3

Position2.4:(Figure 4.16)

The arm and body is on the same line, and the left leg and right leg is 135 degrees.

If the degree of the arm and body are between 170 and 190 degrees, and the left leg and right leg is between 125 and 145 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

Figure 4.16 Position2.4

Position3.1:(Figure 4.17) The right arm and left arm are horizontal, and the left leg and right leg are 135 degrees.

If the degree of the right arm and left arm to the upper body is between 80 and 100 degrees, and the left leg and right leg are between 125 and 145 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

Figure 4.17 Position3.1

Position3.2:(Figure 4.18)

The right arm is vertical to the floor, the left leg and right leg are 135 degrees, the upper body and right leg are 30 degrees, and the right arm and upper body are 180 degrees.

If the degree of the right arm is between 70 and 90 degrees to the floor, the left leg and right leg are 135 degrees, the upper body and right leg are 30 degrees, and the right arm and upper body are 180 degrees, the system will give OK sign, otherwise, it will give ERROR sign.

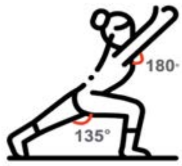


Figure 4.18 Position3.2

Position3.3:(Figure 4.19)

The left arm and right arm are 90 degrees, and the left leg and right leg are 90 degrees too.

If the degree of the left arm and right arm, and the degree of the left leg and right leg are between 80 and 100 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

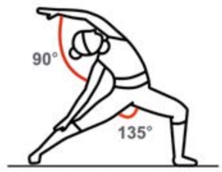


Figure 4.19 Position3.3

Position3.4:(Figure 4.20)

The left arm and right arm is 135 degrees, and the left leg and right leg is 135 degrees too.

If the degree of the left arm and right arm, and the degree of the left leg and right leg are between 125 and 145 degrees, the system will give the OK sign, otherwise, it will give the ERROR sign.

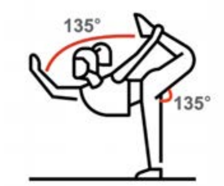


Figure 4.20 Position3.4

Above are the drawing of the positions, but to make sure the participants can clearly know what position they needs to perform, we found according picture with real people performing yoga. The pictures will be shown on the TV while the participants conduct the experiments. For example, the pictures of real person of set 3 is shown below: (Figure 4.21)

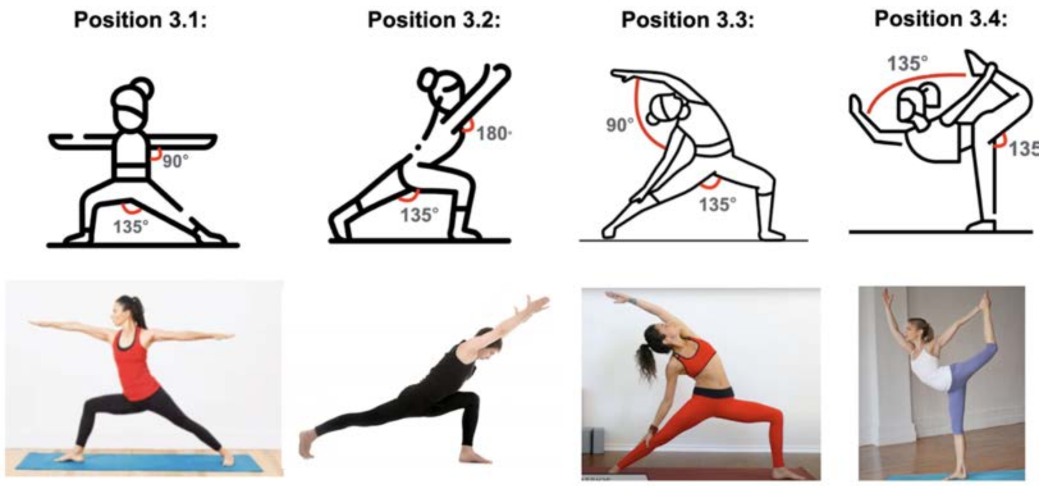


Figure 4.21 Real Person's Yoga Position

Although the prototype is supposed to use camera to detect real time movement, and give automated audio and arrow feedback by comparing user's positions and trainer's positions. At this period, we aren't able to realize the function, what we can do is to record the participant's video and upload to the prototype, get OK or Error instruction. For the audio and arrow feedback, we adopted a semi real-time method by playing slides to indicate the different arrow and audio to correct various body parts. When the body is needing to be adjusted, we show according slide to indicate the direction.

The arrow feedback will be given on-screen, which changes in size to tell the user how much they should lift up or lower their body parts. To give audio feedback, we insert the prerecorded audio into the slide, and give the audio feedback according to the real time body movement of participant. The arrow feedback is also indicated in the slide and which is premade too.

As the audio & arrow feedback is not always the best solution, since some people don't like hearing oral instruction during workouts, the group of "audio & arrow feedback to be given at the same time" is also tested.

The prerecorded voice and premade arrow played in the experiment includes separate sets of orders: Lift up your left leg, lift up your right leg, lift up both leg, lower your left leg, lower your right leg, lower both legs, lift up your left arm, lift up your right arm, lift up both arms, lower your left arm, lower your right arm, lower both arms. The combined set of orders is also included to test the effects of doing workout on leg and arms together, which includes below orders: Lift up your left leg and lift up your right leg; lift up your left leg and lower your right leg; lift up your left leg and lift up your left arm; lift up your left leg and lower your left arm; lift up your left leg and lift up your right arm; lift up your left leg and lower your right arm; lift up your right leg and lift up your left arm; lift up your right leg and lower your left arm; lift up your right leg and lift up your right arm; lift up your right leg and lower your right arm; lift up your left leg and lift up your right arm; lift up your left arm and lower your right arm.

Although the goal the participants need reach is explained before the experiment, the angle won't be instructed in audio feedback or voice feedback as the participants won't be aware of small changes as how big is 30 degrees, or 45 degrees. Also for the big changes, the participants can realize by themselves by giving the instruction as lift up or lower their body parts.

During the experiment, we observe the positions and give according feedback by showing the slides. The control panel of all the slides are shown below: (Figure 4.22)

4.3.3 Group Arrangement

To avoid the bias from muscle memory (participants would have muscle memory if they did same positions for multiple times, which will affect the experiment result for the subsequent positions). We use different participants for the three groups (Figure 4.23) and every group of participants does the three different sets of positions while the difficulty level of the three sets is similar.

For the first experiment, the positions performed in the set of three groups is indicated below:(Figure 4.24)

For the first experiment, the set of three groups is indicated below: Participants start doing the workout with arrow feedback, then receive audio feedback, followed

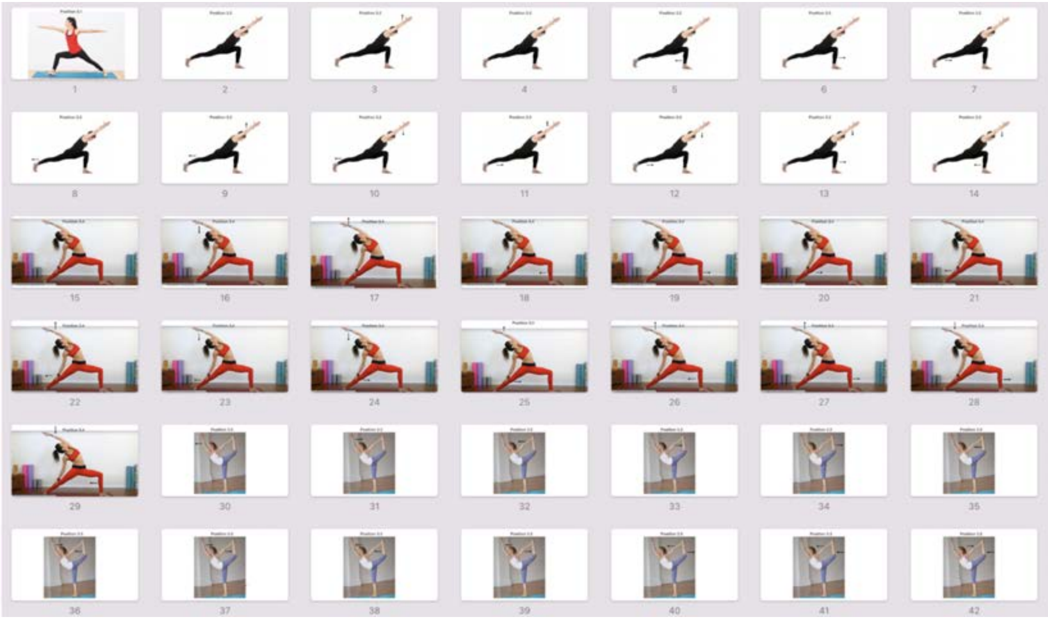


Figure 4.22 Control Panel of the Feedback

by arrow & audio feedback.

The participants will do the workout with feedback for the leg exercises and then for the arm, or receive feedback toward the arm and then toward the leg, which is one set of the positions. Therefore, the participant can compare the three sets of positions and give feedback based on the comparison.

After finishing the positions without feedback, which is the control group, the participants are not asked to answer the questionnaire after the first position is done. Afterwards, after finishing the second workout position with feedback, participants need to answer questionnaire 1 to describe the workout effect between training leg, arm and leg & arm and fill out personal workout experiences with the arrow feedback provided. Also fill out their demographic information.

After finishing the third workout position with feedback, participants need to answer questionnaire 2 to describe the workout effect between training leg, arm and leg arm with the audio feedback provided.

After finishing the fourth workout, which means finishing all the workout, participants are supposed to fill out questionnaire 3 to compare all 4 groups of feedback (audio, arrow, arrow & audio feedback), also separately describe the workout

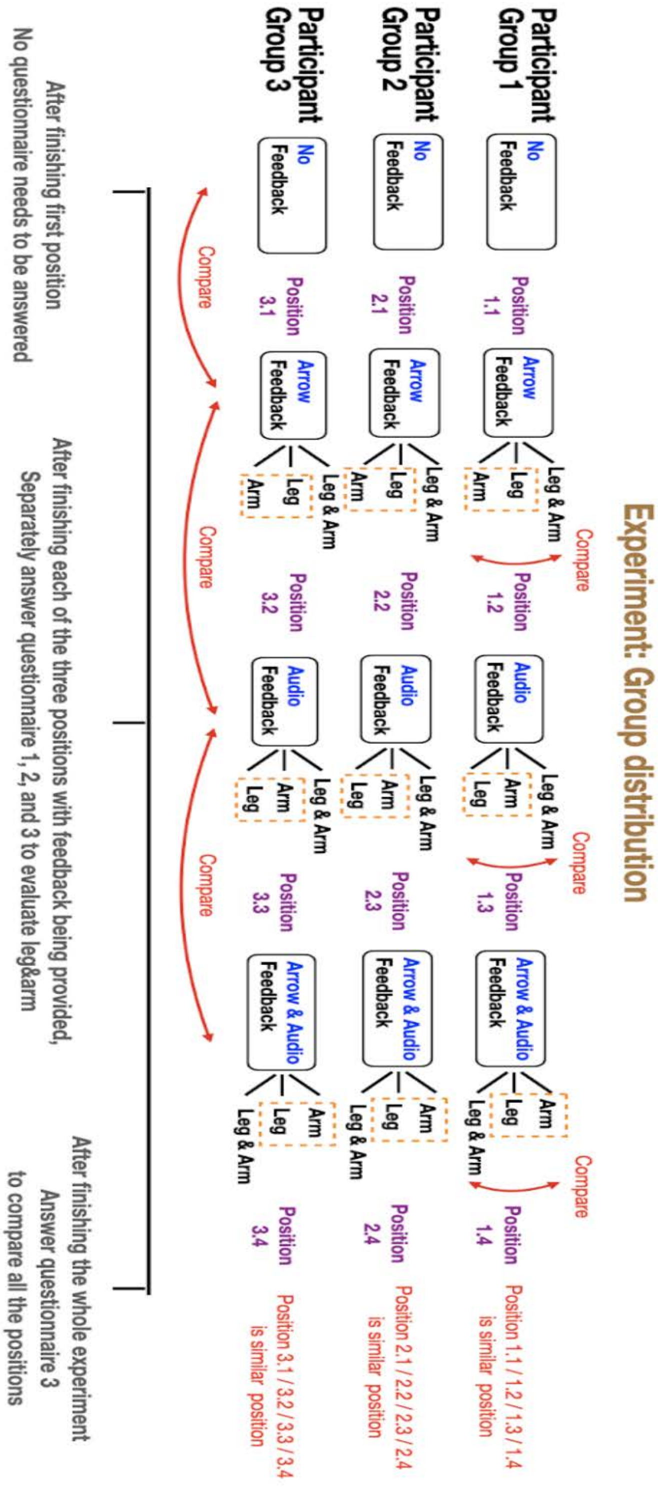


Figure 4.23 Experiment Group Distribution

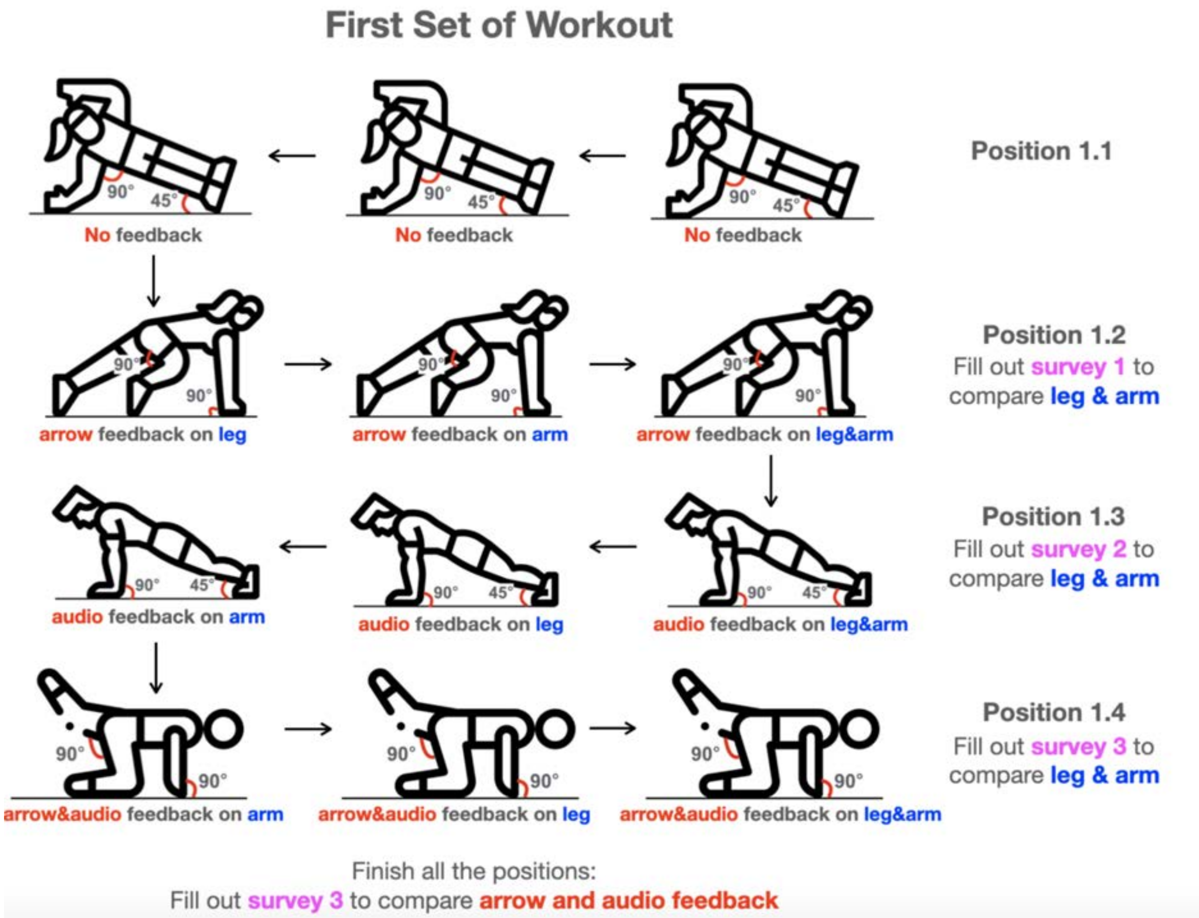


Figure 4.24 Position Performed in the First Set of Workout

effect between training leg, arm and leg arm with arrow and audio feedback provided.

If the sequence of every set of poses (feedback to arm, feedback to leg, and feedback to both arm & leg) stays the same, then the muscle memory toward the same position might cause the bias to see which feedback is the more effective. Therefore, to decrease the effect on muscle memory, the sets of poses should go in a different sequence. Be noted when the feedback is only given to a certain body part, the other body parts have to move too. For example, even if the feedback is only given to the leg, the arm and leg both have to move.

As the participant for each group is the same, there won't be bias from assessment standards based on individual differences. Also, the separate setting (separate leg and arm) and collective setting (train leg and arm together) are tested in three different methods: arrow feedback, audio feedback, and arrow & audio feedback. The result for separate settings and collecting settings is precise.

To help to understand the different workout positions, here we give the diagram of the first set of workouts as an example. The other two sets of workouts are in a similar pattern. Participants are asked to fill out a questionnaire to score leg & arms workout and leg to arm workout after each position, and score the arrow, audio, arrow & audio feedback after all the positions are trained.

4.3.4 Time Distribution

Each position would last for 24 seconds. After starting working out for 6 seconds, the system starts to give audio or arrow feedback. The participant might need to adjust their position multiple times, to lift up or lower body parts. The feedback will be accordingly given, which will last for 12 seconds. At the 12th second, the system stops giving feedback, and the participant has 6 seconds to hold their position.

One set of the position includes 3 different feedback settings: feedback is only given to the leg, or feedback is only given to the arm, or feedback is given to both arm & leg. The feedback given to arm and leg separately should be one group, with a 2 seconds interval afterward separately. The feedback given to both arm and leg together is another group with a 2 seconds interval afterward. The 2 seconds interval time for the participant is to get back to the initial position,

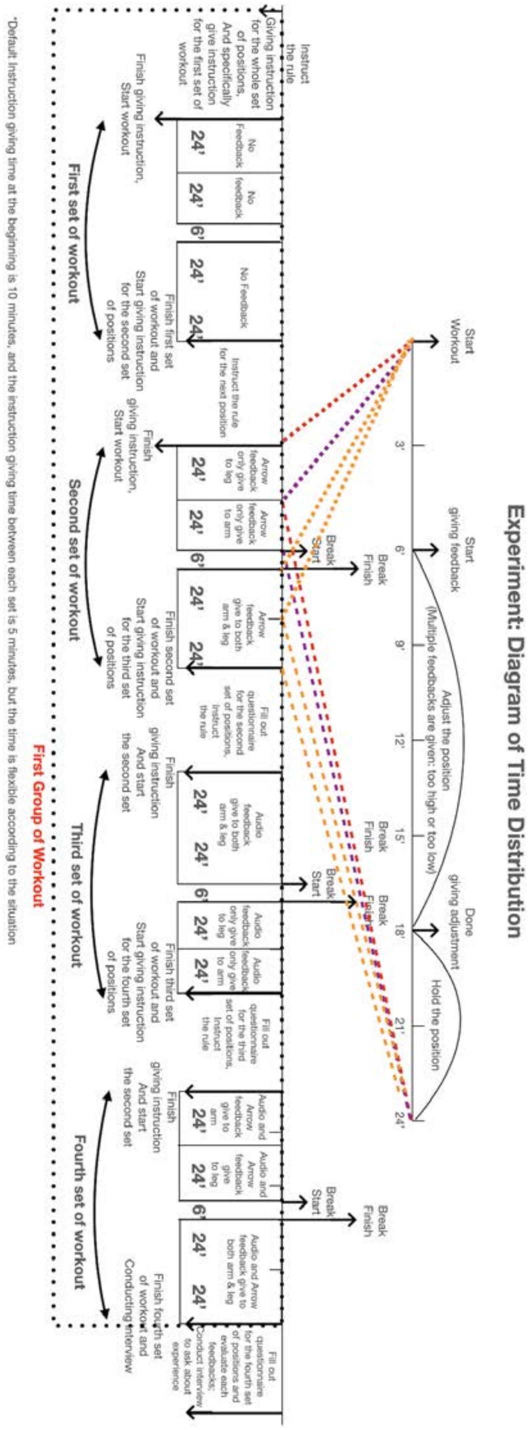


Figure 4.25 Experiment Time Distribution

which will take 102 seconds. Here is the breakdown of the timing: 24 seconds on leg + 24 seconds on arm + interval 6 seconds + 48 seconds (24 seconds each in leg & arm) on arm and leg together = 102 seconds in total for a set of positions.

To ensure the workout time lasts the same between these two groups, the workout time for the collective group (both arm and leg together) is the same as the workout time for the separate group (feedback is given to arm and leg separately). Therefore, the collective group lasts for 48 seconds (24 seconds each in leg & arm).

As each individual spends a different time length to reach their ideal position, participants are allowed to ask for more time to reach the acceptable effect. The extra time spent will also be noted in the report.

Be noted, that the interval instruction time is not fixed as each participant will spend a different length of time understanding the process.

To make sure the participant can get a good understanding of the experiment, prior to starting the workout, I will provide verbal instructions, which includes the angle of positions the participants need to reach to in the experiment, show the picture of the trainer's position, explain the whole process of the workout and the timeline of the workout. Be noted that the instruction given prior to the workout is the most detailed instruction, as the participants need to know the whole process and choose which set of positions they want to do. The instruction given after second and third set of position only explain the key points to keep in mind for the next workout in case the participants forget. In this process, the manuscript is used to give order to avoid any bias caused by the inconsistent order. Also any questions related to the topics will be answered.

After the workout, we conduct a short interview to ask about user's experience, such as "did you have any difficulty on conducting the experiment".

Including the instruction time, training time, time to finish the questionnaire, and break time, the average experiment time should be around 60-90 minutes. (Figure 4.25)

4.4. Evaluation Method

4.4.1 Subjective Way to Evaluate

We ask for subjective opinions through sending questionnaire, which include questions such as “feedback for the experiment”, “Compared to normal workout routine that provides no feedback, I felt my position is closer to trainer’s position during the workout”, “Compared to normal workout routine that provides no feedback, I felt I spent less time to correct my position”, “Among the two workouts you had, which workout is more effective?”, “Please rank the workout you had”, “If you don’t think there are significant difference between the three of them, please write down which feedback and which feedback has similar effectiveness?” “Besides the audio and video feedback, what other feedback do you want to get from workout”. All the answers are subjectively based on participants’ experience.

Also, we set up the recording machine for the experiment. After listening to the recording, we count for how many times of feedback we give to participants and the time spent on adjusting the positions. If the feedback is decreasing, or the time spent on adjusting positions gets shorter compared to the normal workout without receiving any feedback, we count this as “feedback might be helpful”. Vice versa, the feedback doesn’t count as helpful. Also, if there is any comment such as “this position is difficult”or “this is getting easier”, it should be written in the evaluation.

We also receive instant feedback from participants when they conduct experiments, which can be marked from the recording.

At the end of experiments, the interviews are conducted with participants to ask about their feeling and any advice regarding the experiments.

4.4.2 Objective Way to Evaluate

We record the workout video of the experimenter and import it to the python test software we created, and count how many OK / ERROR the user gets for the whole 8 sessions of the workout. If the participants get more OK than failure in each set of workout, the participants’ workout get improved by receiving the feedback. If participant get more OK on the workout which he receives A feedback than the workout which he receives B feedback, the A feedback is more effective than B.

In addition to the total number of OK a participants get in the experiment, the accumulated change of position is also calculated.

The degree of the positions is calculated to see if they reach a higher degree when they try the same position for more times. For example: assuming the degrees of the standard position is 90 degrees, in the first try, the participants were only able to reach a lower degree, such as 60 degrees. In the second try, he reaches a higher angle as the position improves when he gets more practice, which reaches 70 degrees. Although the degree still doesn't pass, the system would recognize this as an improvement on workout effect.

4.5. Result

The questionnaire is evaluated with a 7 point Likert Scale (1=strongly disagree, 4 is neutral/ mixed, 7 is strongly disagree)

4.5.1 Background Study on Participant's Situation

As the below two figure shows, more than 80 percent of users have never done yoga before, this fits the assumption of setting the target group to yoga beginners, (Figure 4.26) while most of them evaluated themselves as relatively good at doing workouts. (Figure 4.27)

Half of the users had experience in using workout videos, from the interview with these people, they showed more preference for using the feedback while using the workout video. (Figure 4.28)

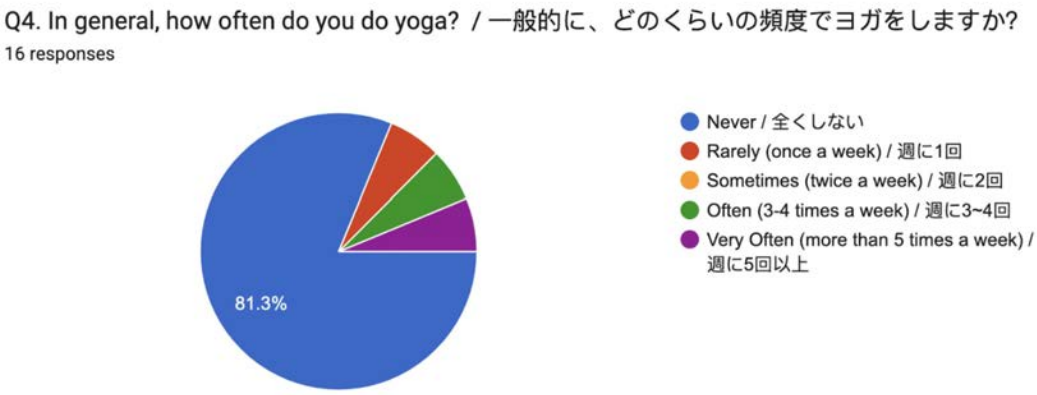


Figure 4.26 Answers for How Often Do You Do Yoga

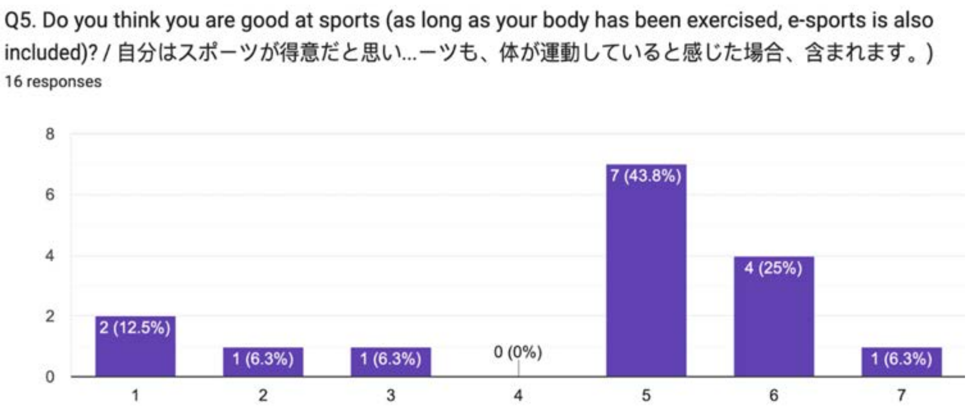


Figure 4.27 Answers for Are You Good at Sports

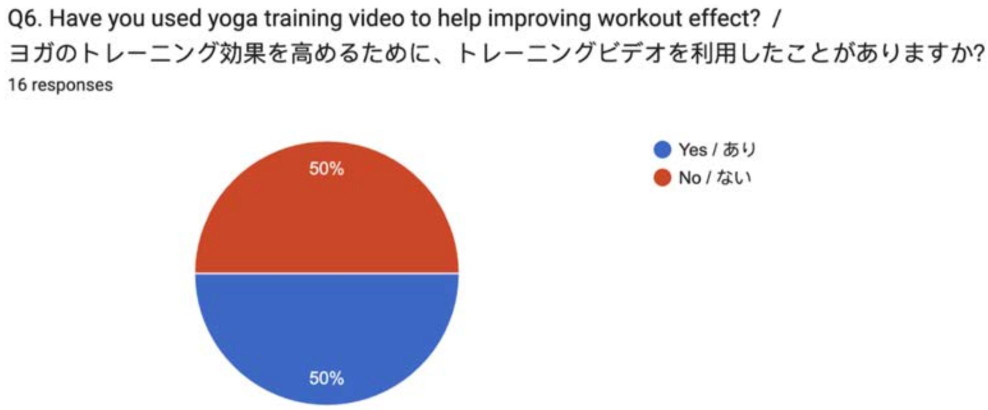


Figure 4.28 Answers for Yoga Video Using Experience

4.5.2 Overview Result for Questionnaire 1: Arrow Feedback

Answers for Participants Getting Closer to Trainer's Position (Arrow Feedback)

For users who get arrow feedback, the graph of their ranking on "if i felt my position is close to trainer's position" is indicated below. (Figure 4.30)

For users who get arrow feedback, if we mark the number below 4 (include 4) as "not getting close enough", and mark the number above 4 as "getting relatively close". Most participants think their position is extremely closer to the trainer's position. The histogram can be seen below to compare the two sets of data: (Figure 4.29)

The number of participants who receive arrow feedback selects "position gets relatively close to trainer's position" is more than the number of participants who receive audio, or arrow & audio feedback together.

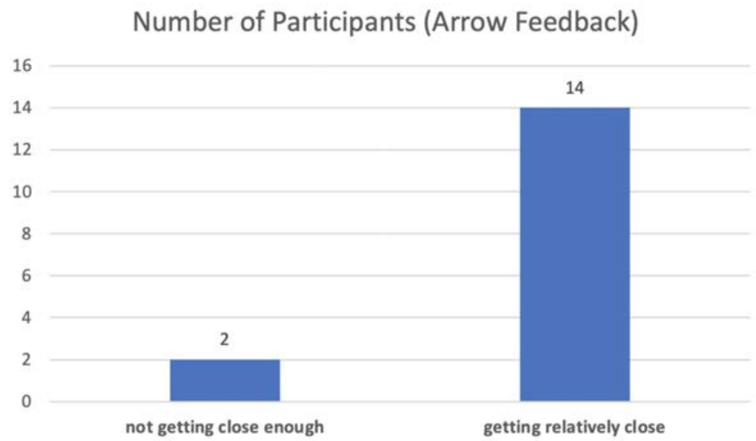


Figure 4.29 Number of Participants Who Feel Their Positions Getting Closer to Trainer's Position (Arrow Feedback)

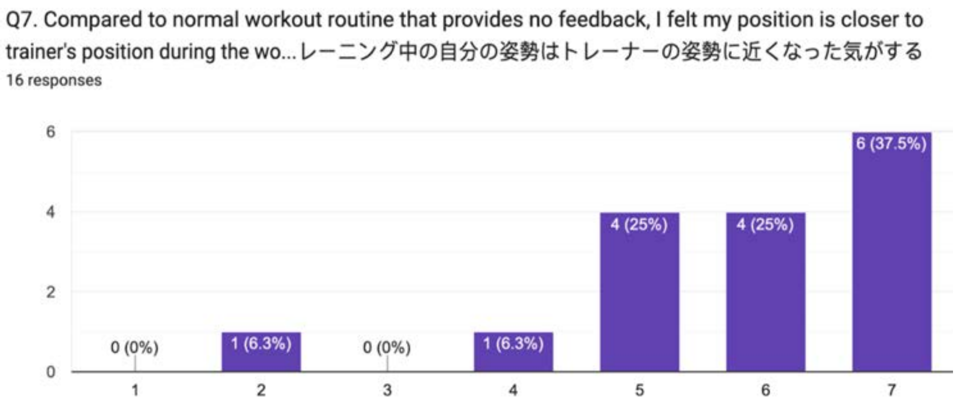


Figure 4.30 Evaluation on Getting Closer to Trainer's Position (Arrow Feedback)

Answers for Participants Spending Less Time on Correcting Position (Arrow Feedback)

For users who get arrow feedback, the graph of their ranking on "if I felt I spent less time on correcting position" is indicated below.(Figure 4.32)

Regarding the time spent on correcting the position, if we mark 4 and below as "adjust position slowly", and mark the number greater than 4 as "adjust position quickly". Less people found themselves adjusting the position quickly compare to the workout without receiving audio feedback,or arrow and audio feedback together, while most people comments that arrow feedback helps them adjust position in a faster speed. The histogram can be seen below to compare the two sets of data: (Figure 4.31)

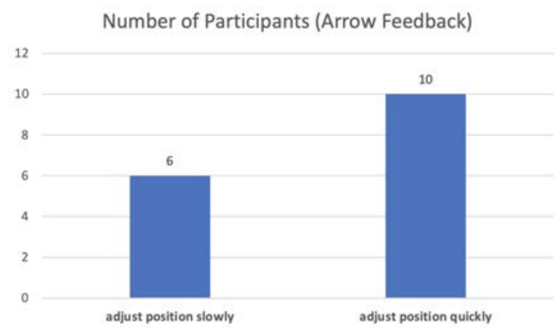


Figure 4.31 Number of Participants Who Feel Their Spent Less Time On Correcting Position (Arrow Feedback)

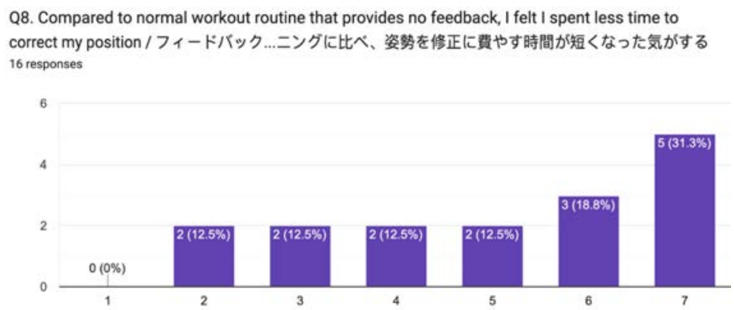


Figure 4.32 Evaluation on Spending Less Time on Correcting Position (Arrow Feedback)

Answers for Comparing Leg and Arm Workout Effect (Arrow Feedback)

For the answers on comparing leg and arm workout effect while receiving arrow feedback, the graph is shown below: (Figure 4.33)

Most people think giving feedback separately to both leg and arm is effective. Giving feedback to both leg and arm is also effective, but not as effective as giving feedback separately.

For the option "giving feedback collectively to both leg and arm is more effective", more people who receive arrow feedback choose it than the people who receive audio feedback, or receive arrow and audio feedback together.

For the option "giving feedback separately to both leg and arm is more effective", the amount of people who choose it while receiving arrow feedback is the same as the amount of people who receive both audio and arrow feedback together. less people who receive arrow feedback choose it than the people who receive audio feedback.

A minority of people don't think it has a significant difference between giving feedback to leg and arm collectively or separately. The number of people who choose it is less than people who receive audio & arrow feedback, but more than the people who receive audio feedback.

Figure 4.33 Comparison for Leg and Arm Workout Effect (Arrow Feedback)

4.5.3 Overview Result for Questionnaire 2: Audio Feedback

Answers for Participants Getting Closer to Trainer's Position (Audio Feedback)

For users who get audio feedback, the graph of their ranking on "if i felt my position is close to trainer's position" is indicated below. (Figure 4.35)

For users who get audio feedback, if we mark the number 4 and below as "not getting close enough", and mark the number greater than 4 as "getting relatively close". The histogram can be seen below to compare the two sets of data: (Figure 4.34)

While for the participants who feels their position are getting relatively close to the trainer, the number of who receive audio feedback is the same for those receive arrow & audio feedback. Comparing the participants who receive arrow feedback feels their position are getting relatively close to the trainer, less participants who receive the arrow feedback feel their position gets very close to the user's position than people.

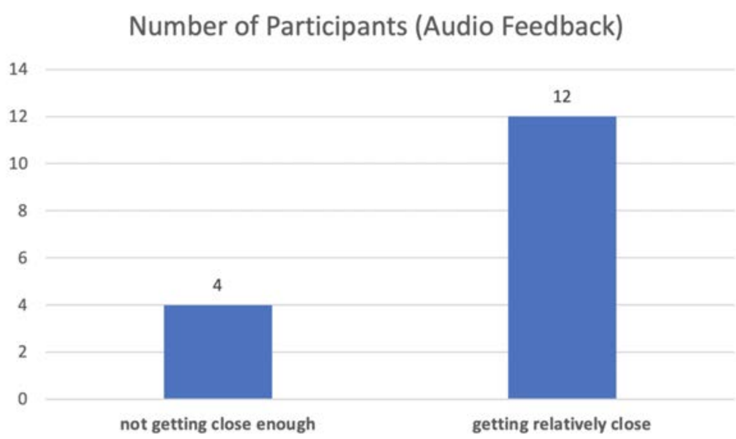


Figure 4.34 Number of Participants Who Feel Their Positions Getting Closer to Trainer's Position (Audio Feedback)

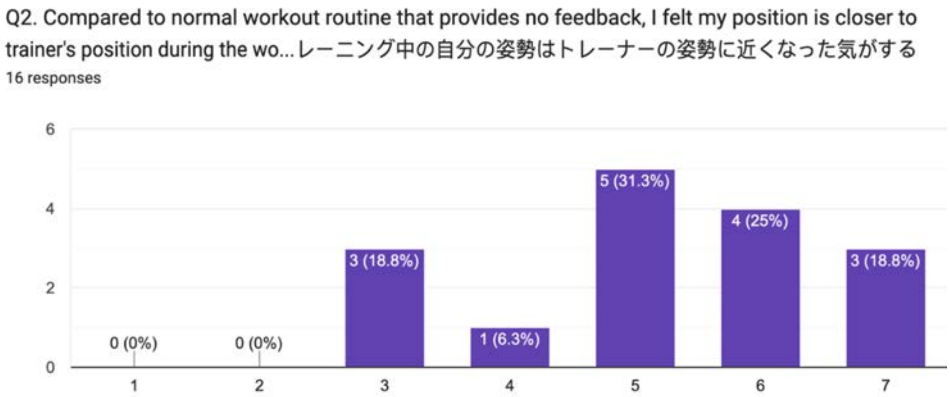


Figure 4.35 Evaluation on Getting closer to Trainer's Position (Audio Feedback)

Answers for Participants Spending Less Time on Correcting Position (Audio Feedback)

For users who get arrow feedback, the graph of their ranking on "if I felt I spent less time on correcting position" is indicated below.(Figure 4.37)

Regarding the time spent on correcting the position, if we mark the number 4 and below as "adjust position slowly", and mark the number greater than 4 as "adjust position quickly". More people found themselves adjusting their position quickly compare to the people who receive audio feedback, while almost the same amount as the participants who receive arrow and audio feedback together. The histogram can be seen below to compare the two sets of data: (Figure 4.36)

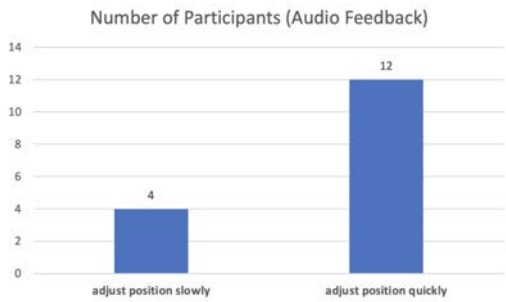


Figure 4.36 Number of Participants Who Feel Their Spent Less Time On Correcting Position (Audio Feedback)

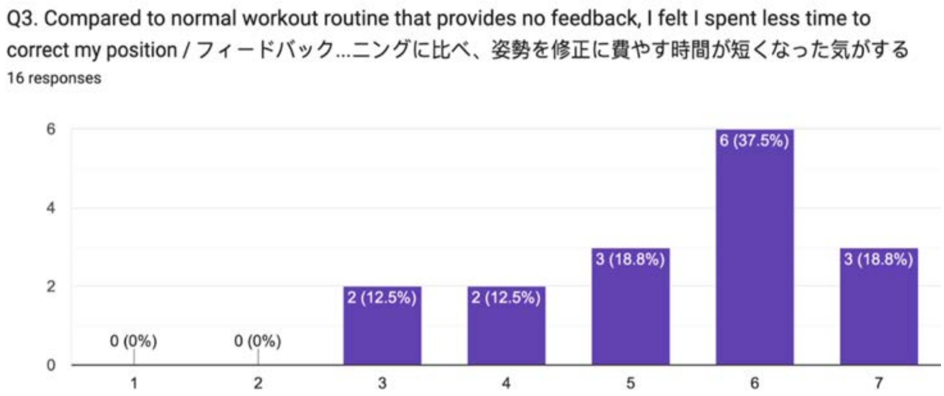


Figure 4.37 Evaluation on Spending Less Time on Correcting Position (Audio Feedback)

Answers for Comparing Leg and Arm Workout Effect (Audio feedback)

For the answers on comparing leg and arm workout effect while receiving audio feedback, the graph is shown below: (Figure 4.38)

Most people think giving feedback separately to both leg and arm is effective, compare to “give feedback collectively to leg and arm” and “no significance between the two option”.

For the option “giving feedback separately to both leg and arm is more effective”, more people who receive audio feedback choose it than the people who receive arrow feedback, while almost same amount as the people who receive both arrow and audio feedback.

For the option “giving feedback collectively to both leg and arm is more effective”, more people who receive audio feedback choose it than the people who receive arrow feedback, while almost same amount as the people who receive both arrow and audio feedback.

Least people choose “there is no significant difference between the two option”.

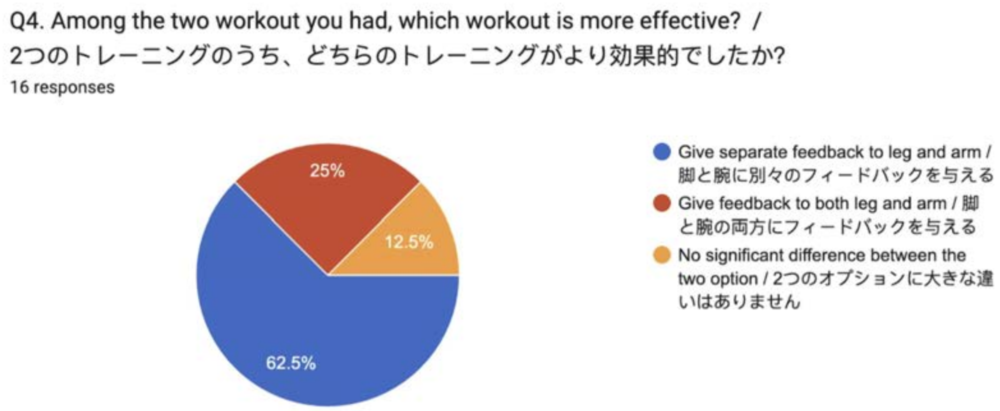


Figure 4.38 Comparison for Leg and Arm Workout Effect (Audio Feedback)

4.5.4 Overview Result for Questionnaire 3: Audio & Arrow Feedback

Answers for Participants Getting Closer to Trainer's Position (Audio & Arrow Feedback)

For users who get audio & arrow Feedback, the graph of their ranking on "if i felt my position is close to trainer's position" is indicated below. (Figure 4.40)

For users who get arrow and audio feedback together, if we mark the number 4 and below as "not getting close enough", and mark the number greater than 4 as "getting relatively close". The histogram can be seen below to compare the two sets of data: (Figure 4.39)

While for the participants who feels their position are getting relatively close to the trainer, the number of who receive audio feedback is the same for those receive arrow & audio feedback. Comparing the participants who receive arrow feedback feels their position are getting relatively close to the trainer, less participants who receive the arrow & arrow feedback feel their position gets very close to the user's position than people.

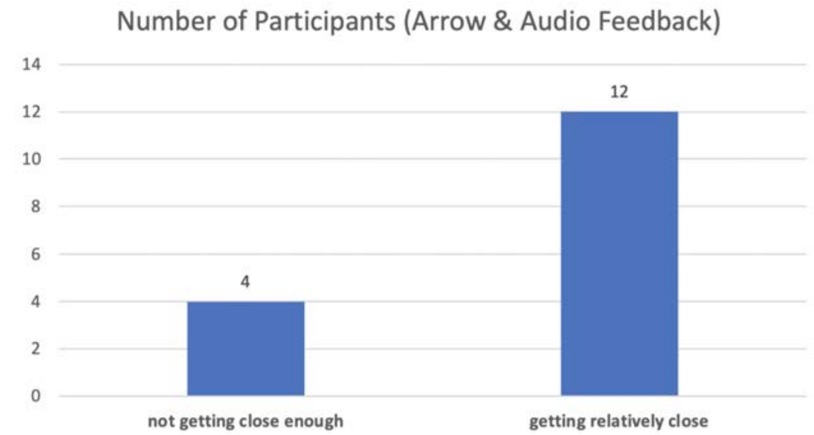


Figure 4.39 Number of Participants Who Feel Their Positions Gets Closer to Trainer's Position (Audio and Arrow Feedback)

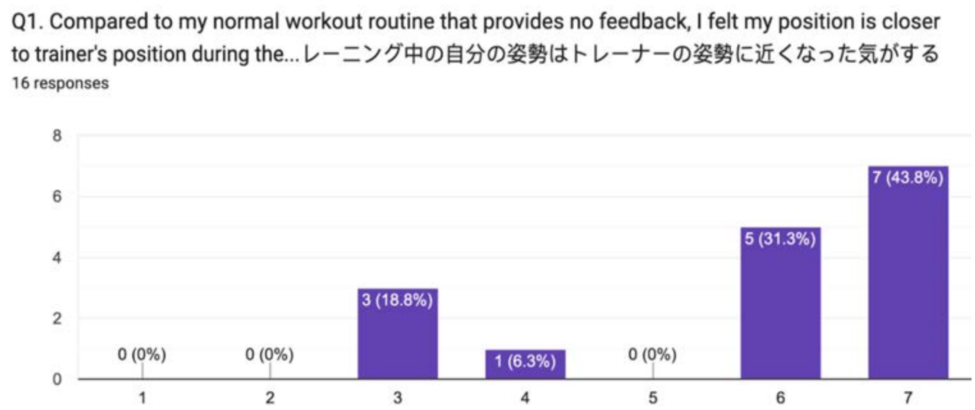


Figure 4.40 Evaluation on Getting Closer to Trainer's Position (Audio and Arrow Feedback)

Spent Less Time on Correcting Position (Audio and Audio Feedback)

For users who get arrow feedback, the graph of their ranking on “if I felt I spent less time on correcting position” is indicated below.(Figure 4.42)

Regarding the time spent on correcting the position, if we mark the number below 4 (include 4) as “adjust position slowly”, and mark the number above 4 as “adjust position quickly”. More people found themselves adjusting their position quickly compare to people who only receive arrow feedback. The number of these participants who choose “getting relatively close” and “getting relatively close” is the same as participants who receive audio feedback.

The histogram can be seen below to compare the two sets of data: (Figure 4.41)

More people found themselves adjusting their position quickly compare to the workout without receiving any feedback, and more people who receive audio feedback adjusted their position compared to the people who receive arrow feedback. The number of participants who choose “getting relatively close” and “getting relatively close” is the same as participants who receive audio feedback. The histogram can be seen below to compare the two sets of data:

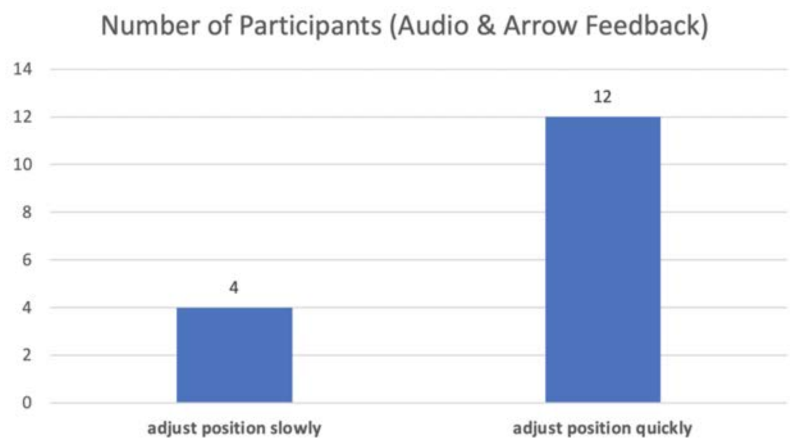


Figure 4.41 Number of Participants Who Feel Their Spent Less Time On Correcting Position (Arrow and Audio Feedback)

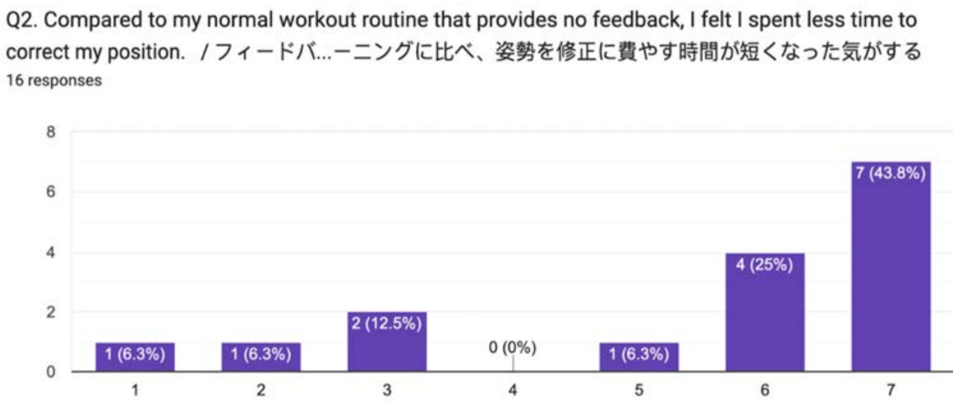


Figure 4.42 Evaluation on Spending Less Time On Correcting Position (Arrow and Audio Feedback)

Answers for Comparing Leg and Arm Workout Effect (Audio & Arrow Feedback)

For the answers on comparing leg and arm workout effect while receiving audio feedback, the graph is shown below: (Figure 4.43)

Most people think giving feedback separately to both leg and arm is effective, compared to the other two options of "giving feedback collectively to leg and arm" and "no significant difference".

For the option "giving feedback separately to both leg and arm is more effective", the least people who receive arrow & audio feedback choose it compared to the people who receive audio feedback, but which is the same amount of people who choose arrow feedback.

For the option "giving feedback collectively to both leg and arm is effective", least people choose it compare to the people who receive arrow feedback, while keep the same amount as the people who receive audio feedback.

Lastly, "no significant difference between the two options", most people who receive both audio and arrow feedback choose it compared to the people who receive either audio or arrow feedback.

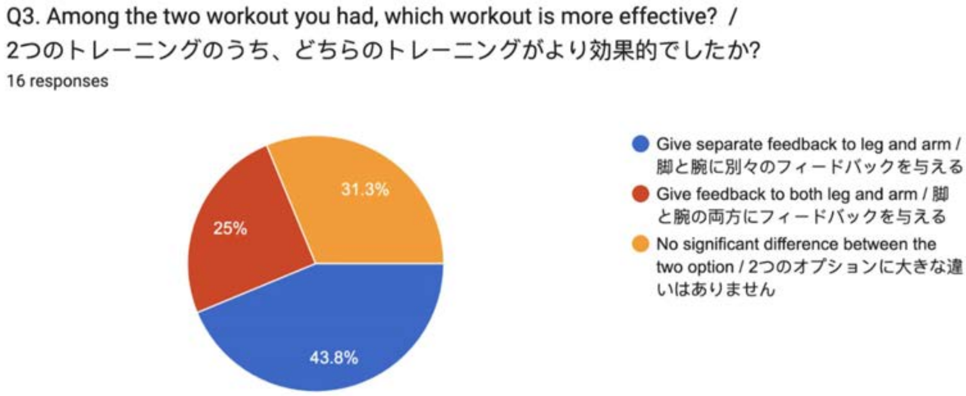


Figure 4.43 Comparison for Leg and Arm Workout Effect (Audio and Arrow Feedback)

Answers for The Effectiveness between Different Type of Feedback

For the below graphs, (Figure 4.44) to compare the effectiveness, it shows that most people think that arrow & audio feedback is the most effective one, the audio feedback is what most people choose for the median effectiveness, the arrow feedback is what most people choose for the least effective option. To compare each feedback, most people think arrow feedback is least effective, the same amount of people think audio feedback can be median and least effective, most people think arrow & audio feedback is the most effective one.

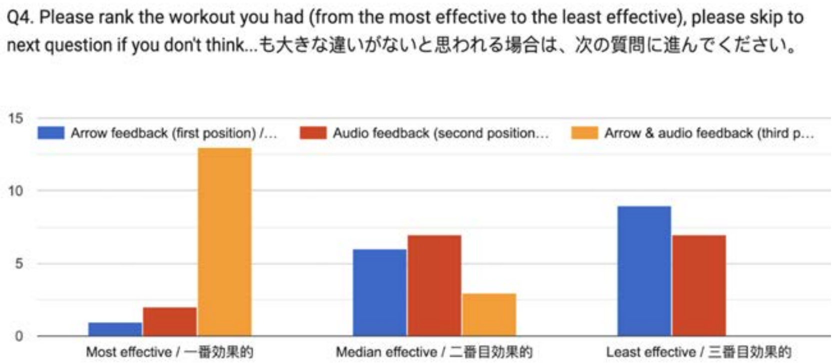


Figure 4.44 The Effectiveness between Arrow, Audio, Arrow and Audio Feedback

Using Mean and SD to Compare the Effectiveness between Different Type of Feedback

For the below graphs, it calculates the mean and SD between the three feedback on the variables "spent less time to correct position" and "get closer to user's position".

The mean value of the amount of people who score is indicated in the Table **??**.

Table 4.1 The Mean Value of the Amount of People Who Score)

	Less Time To Correct Position		Get Closer to Trainer's Position	
	Mean	SD	Mean	SD
Arrow	5.07	1.57	5.69	1.35
Audio	5.38	1.48	5.20	1.51
Audio & Arrow Feedback	5.45	1.97	5.76	1.52

The calculation equation of the mean is:

$$Mean = \frac{(\text{amount of people score 1})^2 + (\text{amount of people score 2})^2 + ... + (\text{amount of people score 7})^2}{(\text{the total amount of participants})} \tag{4.1}$$

The calculation equation of the standard deviation (SD) is:

$$\text{SD} = \sqrt{(\text{amount of people score 1})(\text{score1} - \text{mean})^2 + ... + (\text{amount of people score 7})(\text{score7} - \text{mean})^2} \tag{4.2}$$

For the variable of "spending less time to correct position", the arrow & audio feedback has the best effect, audio feedback is the second best, and the arrow is the worst.

On the other hand, the arrow & audio feedback have highest SD, which indicates the data is more spread out, and the participant's experience has high variety with each individual. The evaluation value isn't clustered near the mean. Audio feedback has the lowest SD, which is a bit lower to the SD value of arrow feedback, while the variation is still higher than 1, so the the observations of the data are still a bit spread out and skewed.

For the variable of "getting close to user's position", the arrow & audio feedback has the best effect, the arrow feedback are the second best, and the audio is the worst.

The arrow & audio feedback, and audio feedback both have high SD, which means the data are spread out a lot, the participants experience may vary between individuals. The arrow feedback has lowest SD, which means data are relatively clustered around the mean, most participants tend to feel the similar way.

The Amount of "OK" Participants for Each Position in Each Group After putting the recorded position into software to test, the amount of OK the participants get are shown below: (Figure 4.45)

	First Group		Second Group		Third Group		
No Feedback	position 1.1		position 2.1		position 3.1		Set 1
	first try	5	first try	5	first try	7	total (first and second try):
	second try	7	second try	9	second try	6	39
	third try	13	third try	14	third try	17	total (third try): 44
Arrow Feedback	position 1.2		position 2.2		position 3.2		Set 2
	first try	9	first try	10	first try	11	total (first and second try):
	second	11	second	14	second	13	68
	third try	24	third try	26	third try	27	total (third try): 77
Audio Feedback	position 1.3		position 2.3		position 3.3		Set 3
	first try	12	first try	12	first try	11	total (first and second try):
	second	15	second	14	second	12	76
	third try	23	third try	22	third try	25	total (third try): 70
Arrow & Audio Feedback	position 1.4		position 2.4		position 3.4		Set 4
	first try	10	first try	10	first try	12	total (first and second try):
	second	15	second	13	second	11	71
	third try	23	third try	24	third try	26	total (third try): 73
							total (first and second try):
							254
							total (all the third try): 264

Figure 4.45 The Amount of "OK" Participants for Each Position in Each Group

One correct position can get one OK. For example: for the audio feedback given to leg, the participant can get OK from it; for the audio feedback given to arm, the participant get OK from it too. For the position which feedback given to both leg and arm, if the participants can get two OK for doing it right. Therefore, if the 16 participants can get every position correct, he will get 16 OK in first try (only get feedback leg), 16 OK in second try (only get feedback arm), and 32 OK in third try (get feedback on both arm and leg).

The positions which give feedback obviously get more "OK" than the positions which gets no feedback, which proves that the feedback helps improve the workout effect.

As for giving collective feedback to leg and arm, which is the total number of the third try. It shows that arrow feedback is the best, audio feedback and arrow & audio feedback is the second and third, while these two values are quite the same. This comes to the same conclusion as we discussed before.

About giving separate feedback to leg and arm, which is the total number of the first and second try. It shows that audio feedback is the best, arrow & audio feedback and arrow feedback is the second and third, while these two values are quite the same. This conclusion also proves what we discussed before (Figure 4.33, Figure 4.38 ,Figure 4.43

The total number of first and second try is a little smaller than the total number of third try, which shows the separate feedback might not as good as collective feedback, while the difference is not too significant.

The accumulative angle adjustments in each participants also verify the same results we discussed before. (Figure 4.30, Figure 4.35 ,Figure 4.40)

The Accumulative Angle Change of Participants

Although some participants don't get an initial Pass from the system, they do improve their angles compared to the trainer as they do the workouts. This is still considered to be a success as the participants get improvement in the workout process. While some of the trials may not follow this, as people may have personal bias toward certain positions, which may require more repetition to train. The general findings of 16 participants and three sets of positions can still be meaningful.

To see the accumulative angle change of the workout, we made line charts to see if the feedback helps improving workout effect.

Here we have three sets of workout positions, so there will be three graphs to compare the effect. The first set is position 1.1, 1.2 and 1.3 (Figure 4.46), the second set is position 2.1, 2.2, and 2.3, (Figure 4.47) the third set is position 3.1, 3.2, and 3.3 (Figure 4.48). As we mentioned before, each set of positions are set to same difficulty level in order to compare and reconfirm the experiment result.

As we have 16 participants, we used the average value of the 16 as the angle number. The number indicated in the y-axis is the difference between the trainer's angle and participants' angle, the smaller the angle difference is, the closer the participants' position is to the trainer's position, and thus the better workout effect it realizes. When the number is smaller than 10 degrees, the participant's position and the trainer is very close and get the PASS sign.

All three graphs show that workout without receiving feedback has the worst effect. In two out of three sets, audio arrow feedback being given together realize the best workout effect at last try. This fits the conclusion we made previously about audio and arrow feedback being the best for their workouts.

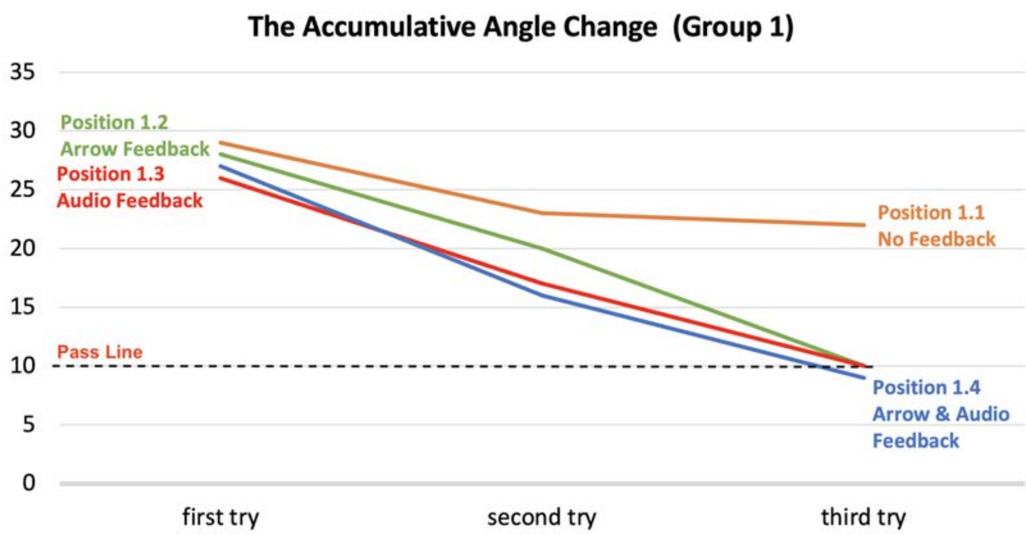


Figure 4.46 The Accumulative Angle Change (Group 1)

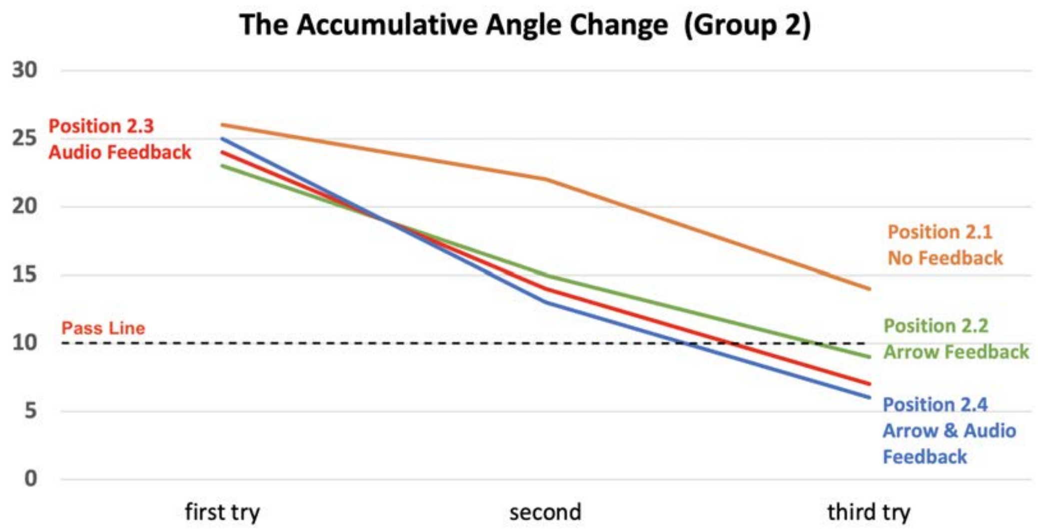


Figure 4.47 The Accumulative Angle Change (Group 2)

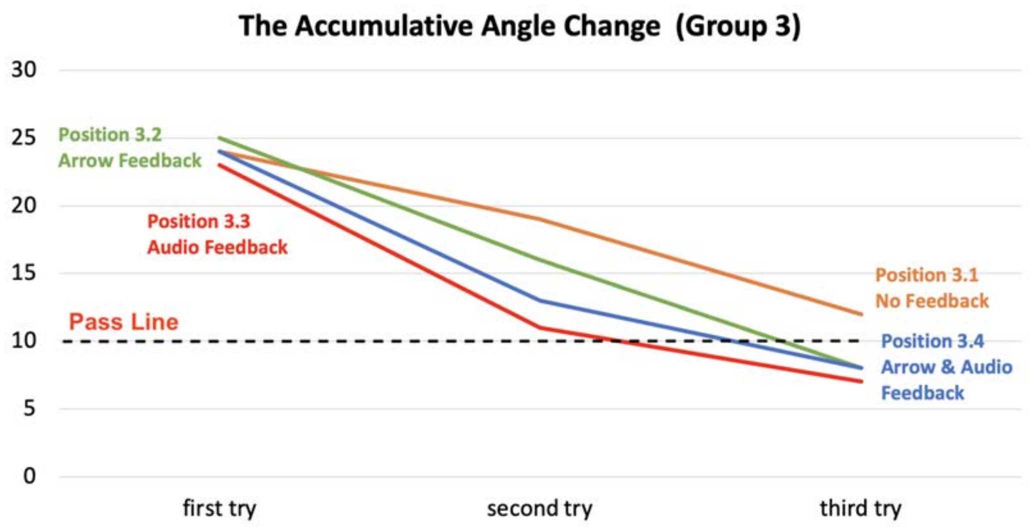


Figure 4.48 The Accumulative Angle Change (Group 3)

Chapter 5
Conclusion and Future Works

5.1. Discussion

5.1.1 Compare Audio & Visual Feedback

Arrow Feedback:

Pro:

Some participants comment that arrow feedback is easier to understand, as it is more intuitive, they only need to look at the screen for one or two seconds to figure out the positions. This is more efficient than audio feedback, as they don't need to wait to have the whole audio sound be played.

Con:

As the visual feedback is shown on the one screen, the participant has to look at the screen, their head as to move, which might affect their body position to be not lined up properly. Also, it is not easy to move the head to look at the screen for certain positions, this can disturb the current balance status. The participants might also get distracted by the arrow feedback and not be able to fully focus on their positions. Therefore, the visual feedback might not get the participants full efficiency of the workout.

Audio Feedback:

Pro:

We received 7 out of 16 participants felt it is easier to adjust their body positions more quickly and accurately when they get audio feedback compared to visual feedback. They also left the comment on audio feedback about "easier to concentrate on", because the participants don't have to look at the screen, and it is much clearer to follow.

Con:

Some people reflected that audio feedback might be disturbing as their don't wish to hear any sound besides the background yoga meditating music.

Audio & Visual Feedback:

Pro:

The audio feedback and visual feedback is given together can collaboratively stimulate the senses to be more focused and easier to understand.

Con:

When the audio and arrow feedback is given simultaneously, some participants might not know which should focus on the most, some would listen to the audio first, which made the arrow feedback become a reconfirming process. If the participants already made the position correct, the arrow feedback is not be helpful and even causes a distraction.

5.1.2 Compare Giving Feedback to Leg and Arm Collectively & Separately

Giving Feedback Collectively:

When it comes to giving collective feedback to both leg and arm, most participants vote for arrow feedback. As seeing different arrows function in multiple body parts is clearer to tell than hearing audio feedback all together at one time, which can cause the distraction as participants would found it hard to focus on which audio feedback should they start first.

Giving Feedback Separately:

Most participants feel that giving separate audio feedback to leg and arm make the difference on the workout effect. The reason can be that participants can be more focused on audio when they hear sound instead of looking at the screen. At this time, giving audio feedback to participants step by step help them focus more on their movement.

5.2. Limitation

In some cases, the participants only need to adjust their arm position as their leg positions are accurate, so the arm & leg feedback given together has no different from arm & leg feedback given separately. This led to the result that participant would choose "no significant difference on getting feedback to leg and arm separately or collectively".

If the position is difficult, they wish to receive arrow and audio feedback separately, as they won't be able to get multiple body joints adjusted at the same time. While if the position is easy, the participants wish to receive multiple feedback together, so they can adjust the feedback more efficiently. Some participants gave feedback that the positions selected in this experiment need to be more difficult, so they can feel the feedback is helpful as they can feel big difference between the initial workout and the workout with feedback being provided.

The subjective survey is affected by some interference factor, such as the participant's mood. Also, in the survey, we only listed "I felt my position is closer to trainer's position" and "I felt I spent less time to correct my position" to evaluate the workout effect, but there might be other factors reflect workout effect which we didn't include in the survey.

In order to look at screen, the trainer has to constantly turn their head to face to the screen, which can effect the workout affect as the head movement will affect the rest of the body, also limit the type of positions user can do.

The sample number is too small, which only include 16 participants, which impact the accuracy of result.

5.3. Conclusion

We proposed a system to collect and compare the pose of workout users and trainers, and then give correction advice according to the comparison result.

The user study showed the poses predicted by the system help improve the workout effect in most cases:

Regarding visual and audio feedback, people who are more responsive to visuals than audio, would prefer to receive arrow feedback. The audio feedback provided might not form positive effects on them. Vice versa, people who are more responsive to audio than visual, would prefer to receive audio feedback. The visual feedback provided might not form positive effects on them.

To have the position closer to the trainer's position, arrow & audio feedback giving together, and only giving arrow feedback can be the most effecting feedback. to To help adjust body in a quick speed, receiving arrow & audio feedback together, and only receiving audio feedback can be the most effective feedback.

As for the giving feedback collectively or separately, people would like to receive feedback to different body parts separately when they get audio feedback. When people receive arrow feedback, they prefer to receive feedback on different body part collectively.

5.4. Future Work

5.4.1 Prototype Improvement

Despite the prototype was not quite finished to reach the effect we want, we are able to take video for the user, and compare the user's video and the trainer's video to give correct or wrong advice.

If the technology allows, the feedback should be automated and given by a machine in real-time, the camera should recognize the positions, and all the feedback are given by precisely comparing the user's position to the trainer's position. Also, the arrow should be given right to the point where the trainer's position is at.

At this case, the remaining question are to be discussed:

Computational delay causes inconvenience. As the device needs some time to collect data and analyze data, although it is desirable to keep the computational delay short, the processing time would still have some delay.

The difference of angle between user and trainer is limited to a certain degree. Here the system set threshold number as 10, it will only give error when the angle different between user and trainer is greater than 10. Any error indicated to be smaller than 10 can't be given ERROR message, but sometimes 10 degrees might be too small or too big to different participants, as every participants has their preference on workout, it can be strict or loose.

Using different cameras might affect the precision of data and multiple devices application. The application of identical camera would realize ideal affect. The standard camera we used in this research is iPhone 11 built-in camera, the other camera can also be used, but would somewhat affect the precision.

Background with a lot noisy color is hard to be precisely recognized by RGB camera. Plain background will detect the position more precisely. The experiments are all conducted under the white background.

Also, OpenPose have problems estimating pose when the ground truth example has non typical poses and upside down examples. In highly crowded images where people are overlapping, the approach tends to merge annotations from different people

Our device only accepts 2d data and only supports almost front-angled positions, changing the position to side profiles or twisting the angle would affect the precision of the result. This is because we assume users use RGB cameras. If we can capture 3D data, we could easily extend our system to 3D data. However, obtaining 3D pose data from an image can be difficult, and this could also be studied in future work.

The method is only limited to the comparing photo and video, when the movement reach to the same position as the photo shows, the system will give OK signal. This limited the prototype can only be used in static yoga positions, other than dynamic positions and other gymnastics training.

5.4.2 More Function

The technology will be considered in application on smartphones, head-up displays, and other devices. This can be a module called history record, which gives advice based on the previous workouts. A customized workout plan can also be offered, which arranges different workout courses and give instruction on how frequently the user should do it. This is based on the pre-workout questionnaire about the user's needs: "are you doing a workout to lose weights/gain muscles/maintain health", and how many years of workout experience do the user have. The user can also check the historical score they can see their workout time, and workout score every week and every month. For each position the participants are going to work on, the instruction of which body parts are going to be exercised, and which body parts might feel painful afterward will be detailed given.

Also, the app also gives encouragement to stimulate users to do workouts, which is mentioned by two experimenters in the interview after the experiment. For example, when the participants get improvement compared to the last time, the screen has a "flower drop on the screen", and when the participants did the position perfectly, the screen shows a "smiling face" or applauding sound. This function can make the workout experience more enjoyable.

The app also can give warnings on the positions in which is easy to get hurt. If the participants raise their leg too high and which can hurt themselves, the system will give voice feedback to remind them "don't lift up your leg too high, you will hurt yourself". In this way, the user will be protected with the warning.

The other function such as when to inhale and exhale can also be given by audio, so the beginner would know how to control their breath.

To remind the experimenter when to finish the workout, the app also gives the count-down sound for the last 5 seconds, so the user has the motivation to hold the positions when they found it difficult to hold in the last few moments. The tool such as wearable sensors can also be combined to add to the effects. The wearable sensors can give feedback by sending vibrations when the position is not correct. This way the participants can clearly tell when their positions are not precisely correct.

Stuff like lasers can also be given to participants to tell them what is the right position to reach, and the standard line would be easy to tell as the laser line can be seen right around the body.

The screen can also be set in 360 degrees, so the arrow feedback can be seen from any direction without getting the head turned. This way the user's position won't be affected by the fixed location of the screen.

The participants can also select if the feedback should work on separate body parts or multiple body parts together according to their preference. The visual and audio effects are also open to switching based on their favorites.

In the future, the app can be used in comparing user's video with trainer's video, which is not limited to the static picture. This allows the position to be dynamic, which can be applied in not only yoga, but also dance, weight lifting, and etc.

References

[1] Alex Polish. Hey, powerlifters – don't be embarrassed to continue group fitness after quarantine. In *Barbend*, 2020. URL: https://barbend.com/powerlifting-group-fitness/.

[2] Nicholas Rizzo. 1/3 gym members won't return after vaccine (11k surveyed). ResearchCenter, 2021. URL: https://runrepeat.com/gyms-post-covid-vaccine.

[3] Jackie Davalos. How covid-19 has permanently changed the fitness industry. Bloomberg, 2020. URL: https://www.bloomberg.com/news/articles/2021-01-19/fitness-industry-may-never-return-to-its-old-ways-after-covid-19.

[4] RD CPT Katey Davidson, MScFN. 9 fitness trends to expect this year. 2021. URL: https://www.healthline.com/nutrition/fitness-trends.

[5] Carmen Ang. Miscthe growth of home fitness apps. 2020. URL: https://www.visualcapitalist.com/the-growth-of-home-fitness-apps-2020/.

[6] Phil Blechman. Covid-19 and american gyms — revenue down 15.6 percent, memberships could fall 40 2021. URL: https://barbend.com/covid-19-american-gyms/.

[7] BRAD SHAW. Rise in exercise related injuries as people workout at home. FITNESS EDITOR, 2020. URL: https://sustainhealth.fit/lifestyle/rise-in-exercise-related-injuries-as-britons-workout-at-home/.

[8] Coral Murphy Marcos. Coronavirus bingeing: Youtube trends show spike in every-day task tutorials. 2020. URL: https://www.usatoday.com/story/tech/2020/05/01/coronavirus-youtube-task-tutorials/3044564001/.

[9] Jun Rekimoto Katsuhito Sasaki, Keisuke Shiro. Exemposer: Predicting poses of experts as examples for beginners in climbing using a neural network. pages 1–9. AHs '20: Proceedings of the Augmented Humans International Conference, 2020.

[10] Jennifer Lee Hicks Apoorva Rajagopal Lukasz Kidziński, Bryan Yang. Deep neural networks enable quantitative movement analysis using single-camera videos. page 11(1):4054. Nature Communications, 2020. URL: `https://www.researchgate.net/publication/343632899_Deep_neural_networks_enable_quantitative_movement_analysis_using_single-camera_videos`, `doi:10.1038/s41467-020-17807-z`.

[11] Elisa D. Mekler Anna Lisa Martin-Niedecken. The exercube: Participatory design of an immersive fitness game environment. In *Serious Games, Proceedings of the 4th Joint International Conference, JCSG 2018*, pages 263–275. IEEE, 2018. URL: `https://www.researchgate.net/publication/328313566_The_ExerCube_Participatory_Design_of_an_Immersive_Fitness_Game_Environment`, `doi:10.1007/978-3-030-02762-9_28`.

[12] Jun Rekimoto Katsuhito Sasaki, Keisuke Shiro. Exemposer: Predicting poses of experts as examples for beginners in climbing using a neural network. 2020. URL: `https://dl.acm.org/doi/10.1145/3384657.3384788`.

[13] Jiyoung ParkJ Samyeul Noh, Sang-baek Lee. Intelligent interaction system capable of autonomous mobility for at-home workouts. pages 653–657. IEEE, 2019. URL: `https://ieeexplore.ieee.org/document/8939887`, `doi:10.1109/ICTC46691.2019.8939887`.

[14] Samyeul Noh Jiyoung Park. Utilization of mobile robot at interaction systems for at-home workouts: Provision of autonomous mobility in indoor environments. pages 1160–1163. IEEE, 2019. URL: `https://ieeexplore.ieee.org/document/8971505`, `doi:10.23919/ICCAS47443.2019.8971505`.

[15] Hideki Koike Atsuki Ikeda, Dong-Hyun Hwang. Real-time visual feedback for golf training using virtual shadow. page 445–448. Proceedings of the

2018 ACM International Conference on Interactive Surfaces and Spaces, 2018. URL: https://dl.acm.org/doi/abs/10.1145/3279778.3279927, doi:10.1145/3279778.3279927.

[16] Tomas Simon Shih-En Wei Yaser Sheikh Zhe Cao, Gines Hidalgo. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. pages 1–14. IEEE, 2019. URL: https://arxiv.org/pdf/1812.08008.pdf5, doi:5.

[17] Arthur Zuckerman . 51 gym membership statistics: 2020/2021 data, trends predictions. CompareCamp, 2020. URL: https://comparecamp.com/gym-membership-statistics/.

[18] Mrinal Singh Walia. A comprehensive guide on human pose estimation. In *Data Science Blogathon*, 2022. URL: https://www.analyticsvidhya.com/blog/2022/01/a-comprehensive-guide-on-human-pose-estimation/.

Appendices

A. Experiment Procedure

A.1 Explanation of Experiment

Participant Number __

Explanation 説明書

Handling of personal information and data 個人情報とデータの取り扱い

The acquired data and personal information will not be used for any purpose other than research purposes. Research personal information will not be disclosed when making a research announcement. However, age and gender may be recorded but not compulsory.

取得したデータや個人情報は、研究目的以外には使用しません。研究発表する際も個人情報は公開致しません。ただし、年齢と性別に関しては記録する場合がありますが、必須ではありません。

About the rights of experiment participants 実験参加者の権利について

You are free to decide to participate in this study. Even after you have agreed once, you can revoke your consent. There is no disadvantage due to it. The data and analysis results obtained from the relevant participants will be discarded and will not be used for further research. However, the results were announced at the time of the destruction request. Please note that if there is, the data cannot be destroyed.

この研究への参加は自由意志で決定してください。一度同意した後でも、同意を破棄することができます。それによる不利益はありません。該当参加者から取得したデータや解析結果は破棄され、それ以降の研究には使用しません。ただし、破棄要求時点で公表済みの結果がある場合、データを破棄することはできませんのでご了承ください。

About intellectual property rights 知的財産権について

Intellectual property rights may arise as research progresses, but the attribution of intellectual property rights in this research is determined in consultation with researchers, Keio University, or collaborators outside Keio Gijuku, and experiments are conducted. It does not belong to the participants.

研究の進展により、知的財産権が生ずる可能性がありますが、本研究における知的財産権の帰属は、研究者または慶應義塾、あるいは慶應義塾外の共同研究者と協議の上決定され、実験参加者に帰属することはありません。

About the experiment 実験について

Participants are asked to do 8 sets of positions (one positions needs to be repeated for 3 times to make it one set, participants can ask to do workout more times to reach to the acceptable affect). There will be visual and audio feedback be provided, and frequency of the feedback will be changed in different groups.
The standard experiment time is around 40-60 minutes.

参加者は 3 セットのヨガ動作をします。実験では、視覚と聴覚のフィードバックが行われ、フィードバックの順番が変わります。

参加者は 1 つの動作を 3 回繰り返すと 1 セットとなり、3 セット行います。ただし、参加者は求める効果に達するまで何回でもポジションを繰り返すことが可能です。標準的な実験時間は約 60~90 分です。

Contact information 問い合わせ連絡先

If you have any inquiries regarding this research, please contact the following.
なお、本研究に関する問い合わせを行う場合は以下連絡先へご連絡ください。

〒223-8526　横浜市港北区日吉 4-1-1 慶應義塾大学 日吉学生部 メディアデザイン研究科

Creative Industry Project: Xiao Qiao 肖僑

xioxio@keio.jp
TEL: 080-7603-7778

以上、何かご不明な点がありましたら遠慮なくお尋ねください。本研究へのご理解とご協力に深く感謝いたします。

A.2 Agreement About Photography

About Photography
撮影について

Photos of the venue, performance and audience members will be taken and audio and video will be recorded before, during and after the study. The recorded material may contain aspects of the performance, the venue as well as members of the audience. The material will be used for documentation purposes, for reporting to Keio University Graduate School of Media Design and it may be submitted to academic conferences and/or published otherwise. Photos, videos and audio files depicting audience members during the performance may be posted online or publicized otherwise by Keio University, academic societies and other institutions.

実験参加中の模様を撮影した動画、写真を、慶應義塾大学大学院メディアデザイン研究科の実施報告、学会投稿に使用させていただく場合があります。その際、お客さまの様子が記録された写真、動画が、慶應義塾大学大学院メディアデザイン研究科の実施報告、学会投稿による WEB などへの露出/掲載をさせていただく場合があります。

Agreement 同意書

I understand the above explanation and agree to participate in this study.
私は、以上の説明を理解し、本研究に参加することに同意します。

Year/年 Month/月 Day/日

PrintName:________________________________

Signature 氏名: ________________________________

A.3 First Questionaire after Receiving Arrow Feedback

Workout Survey 1 (after receiving arrow feedback)

Research Goal:
By doing individual positions (doing one position at a time, such as: doing leg first, then do the arms) or group positions (doing multiple positions at a time, such as: doing leg and arm at the same time) and giving different feedback (visual feedback and audio feedback), what is the best feedback and how much feedback is given at one time can help workout users to realize the best workout effect?

Experiment Content:
Participants are asked to do 8 sets of positions (one positions needs to be repeated for 3 times to make it one set, participants can ask to do workout more times to reach to the acceptable affect). There will be visual and audio feedback be provided, and frequency of the feedback will be changed in different groups.
The standard experiment time is around 40-60 minutes.

研究目的:
一つのフィードバック（一回のトレーニングに身体の一部にフィードバックを提供します。例えば、最初は脚に、次は腕に。）または複数のフィードバック（一度に複数のポジションを行います。例えば、脚と腕を同時にフィードバックを提供します。）一度にどれだけのフィードバックを与えることがユーザーに最高のワークアウト効果を実現させるか?
また、視覚フィードバックと音声フィードバックの中で、どのフィードバックを与えるのが最も効果を高められるのか?

実験内容:
参加者は3セットのヨガ動作をします。実験では、視覚と聴覚のフィードバックが行われ、フィードバックの順番が変わります。
参加者は1つの動作を3回繰り返すと1セットとなり、3セット行います。ただし、参加者は求める効果に達するまで何回でもポジションを繰り返すことが可能です。標準的な実験時間は約60～90分です。

Experiment number / 実験番号

Your answer

Q1. Name / 名前 *

Your answer

Q2. Age / 年齢 *

Your answer

Q3. Gender / 性別 *

- ○ Male / 男性
- ○ Female / 女性
- ○ Prefer not to say / その他
- ○ Other: ____

Q4. In general, how often do you do yoga? / 一般的に、どのくらいの頻度でヨガをしますか? *

- ◯ Never / 全くしない
- ◯ Rarely (once a week) / 週に1回
- ◯ Sometimes (twice a week) / 週に2回
- ◯ Often (3-4 times a week) / 週に3~4回
- ◯ Very Often (more than 5 times a week) / 週に5回以上

Q5. Do you think you are good at sports (as long as your body has been exercised, e-sports is also included)? / 自分はスポーツが得意だと思いますか? (eスポーツも、体が運動していると感じた場合、含まれます。) *

	1	2	3	4	5	6	7	
Strongly Disagree / 非常にそう思わない	◯	◯	◯	◯	◯	◯	◯	Strongly Agree / 非常にそう思う

Q6. Have you used yoga training video to help improving workout effect? / ヨガのトレーニング効果を高めるために、トレーニングビデオを利用したことがありますか? *

- ◯ Yes / あり
- ◯ No / ない

Please use the scale below by circling a number between 1 and 7 to indicate how well or poorly each statement corresponds to you. / 以下の記述について、1-7 当てはまるものを選択してください。

1--------------2--------------3--------------4--------------5-------------6-------------7

Disagree Strongly Neutral/ Mixed Agree Strongly

Q7. Compared to normal workout routine that provides no feedback, I felt my position is closer to trainer's position during the workout / フィードバックが無いトレーニングに比べ、トレーニング中の自分の姿勢はトレーナーの姿勢に近くなった気がする *

	1	2	3	4	5	6	7	
Strongly Disagree / 非常にそう思わない	○	○	○	○	○	○	○	Strongly Agree / 非常にそう思う

Q8. Compared to normal workout routine that provides no feedback, I felt I spent less time to correct my position / フィードバックがないトレーニングに比べ、姿勢を修正に費やす時間が短くなった気がする *

	1	2	3	4	5	6	7	
Strongly Disagree / 非常にそう思わない	○	○	○	○	○	○	○	Strongly Agree / 非常にそう思う

Q9. Among the two workout you had, which workout is more effective? / 2つのトレーニングのうち、どちらのトレーニングがより効果的ですか? *

- ◯ Give separate feedback to leg and arm / 脚と腕に別々のフィードバックを与える
- ◯ Give feedback to both leg and arm / 脚と腕の両方にフィードバックを与える
- ◯ No significant difference between the two option / 2つのオプションに大きな違いはありません
- ◯ Other: ______

Q10. Please write down any feedback toward the experiment / 実験に対するご意見・ご感想をお聞かせください

Your answer ______

A.4 Second Questionaire after Receiving Audio Feedback

Workout Survey 2 (after receiving audio feedback)

Research Goal:
By doing individual positions (doing one position at a time, such as: doing leg first, then do the arms) or group positions (doing multiple positions at a time, such as: doing leg and arm at the same time) and giving different feedback (visual feedback and audio feedback), what is the best feedback and how much feedback is given at one time can help workout users to realize the best workout effect?

Experiment Content:
Participants are asked to do 8 sets of positions (one positions needs to be repeated for 3 times to make it one set, participants can ask to do workout more times to reach to the acceptable affect). There will be visual and audio feedback be provided, and frequency of the feedback will be changed in different groups.
The standard experiment time is around 40-60 minutes.

研究目的:
一つのフィードバック（一回のトレーニングに身体の一部にフィードバックを提供します。例えば、最初は脚に、次は腕に。）または複数のフィードバック（一度に複数のポジションを行います。例えば、脚と腕を同時にフィードバックを提供します。）一度にどれだけのフィードバックを与えることがユーザーに最高のワークアウト効果を実現させるか?
また、視覚フィードバックと音声フィードバックの中で、どのフィードバックを与えるのが最も効果を高められるのか?

実験内容:
参加者は3セットのヨガ動作をします。実験では、視覚と聴覚のフィードバックが行われ、フィードバックの順番が変わります。
参加者は1つの動作を3回繰り返すと1セットとなり、3セット行います。ただし、参加者は求める効果に達するまで何回でもポジションを繰り返すことが可能です。標準的な実験時間は約60～90分です。

christinaxiao.scut@gmail.com (not shared) Switch account

* Required

Experiment number / 実験番号

Your answer

Q1. Name / 名前 *

Your answer

Please use the scale below by circling a number between 1 and 7 to indicate how well or poorly each statement corresponds to you. / 以下の記述について、1-7 当てはまるものを選択してください。

1--------------2--------------3--------------4--------------5-------------6-------------7

Disagree Strongly Neutral/ Mixed Agree Strongly

Q2. Compared to normal workout routine that provides no feedback, I felt my position is closer to trainer's position during the workout / フィードバックが無いトレーニングに比べ、トレーニング中の自分の姿勢はトレーナーの姿勢に近くなった気がする *

1 2 3 4 5 6 7

Strongly Disagree / 非常にそう思わない ○ ○ ○ ○ ○ ○ ○ Strongly Agree / 非常にそう思う

Q3. Compared to normal workout routine that provides no feedback, I felt I spent less time to correct my position / フィードバックがないトレーニングに比べ、姿勢を修正に費やす時間が短くなった気がする *

1 2 3 4 5 6 7

Strongly Disagree / 非常にそう思わない ○ ○ ○ ○ ○ ○ ○ Strongly Agree / 非常にそう思う

Q4. Among the two workout you had, which workout is more effective? / 2つのトレーニングのうち、どちらのトレーニングがより効果的でしたか? *

○ Give separate feedback to leg and arm / 脚と腕に別々のフィードバックを与える

○ Give feedback to both leg and arm / 脚と腕の両方にフィードバックを与える

○ No significant difference between the two option / 2つのオプションに大きな違いはありません

○ Other:

Q5. Please write down any feedback toward the experiment / 実験に対するご意見・ご感想をお聞かせください

A.5 Third Questionaire after Receiving Audio & Visual Feedback

Workout Survey 3 (final, after receiving audio and arrow feedback)

Research Goal:
By doing individual positions (doing one position at a time, such as: doing leg first, then do the arms) or group positions (doing multiple positions at a time, such as: doing leg and arm at the same time) and giving different feedback (visual feedback and audio feedback), what is the best feedback and how much feedback is given at one time can help workout users to realize the best workout effect?

Experiment Content:
Participants are asked to do 8 sets of positions (one positions needs to be repeated for 3 times to make it one set, participants can ask to do workout more times to reach to the acceptable affect). There will be visual and audio feedback be provided, and frequency of the feedback will be changed in different groups.
The standard experiment time is around 40-60 minutes.

研究目的:
一つのフィードバック（一回のトレーニングに身体の一部にフィードバックを提供します。例えば、最初は脚に、次は腕に。）または複数のフィードバック（一度に複数のポジションを行います。例えば、脚と腕を同時にフィードバックを提供します。）一度にどれだけのフィードバックを与えることがユーザーに最高のワークアウト効果を実現させるか?
また、視覚フィードバックと音声フィードバックの中で、どのフィードバックを与えるのが最も効果を高められるのか?

実験内容:
参加者は3セットのヨガ動作をします。実験では、視覚と聴覚のフィードバックが行われ、フィードバックの順番が変わります。
参加者は1つの動作を3回繰り返すと1セットとなり、3セット行います。ただし、参加者は求める効果に達するまで何回でもポジションを繰り返すことが可能です。標準的な実験時間は約60～90分です。

christinaxiao.scut@gmail.com (not shared) Switch account

* Required

set number / 実験番号

Your answer

Please use the scale below by circling a number between 1 and 7 to indicate how well or poorly each statement corresponds to you.
以下の記述について、1-7 当てはまるものを選択してください。

1--------------2--------------3--------------4--------------5--------------6--------------7

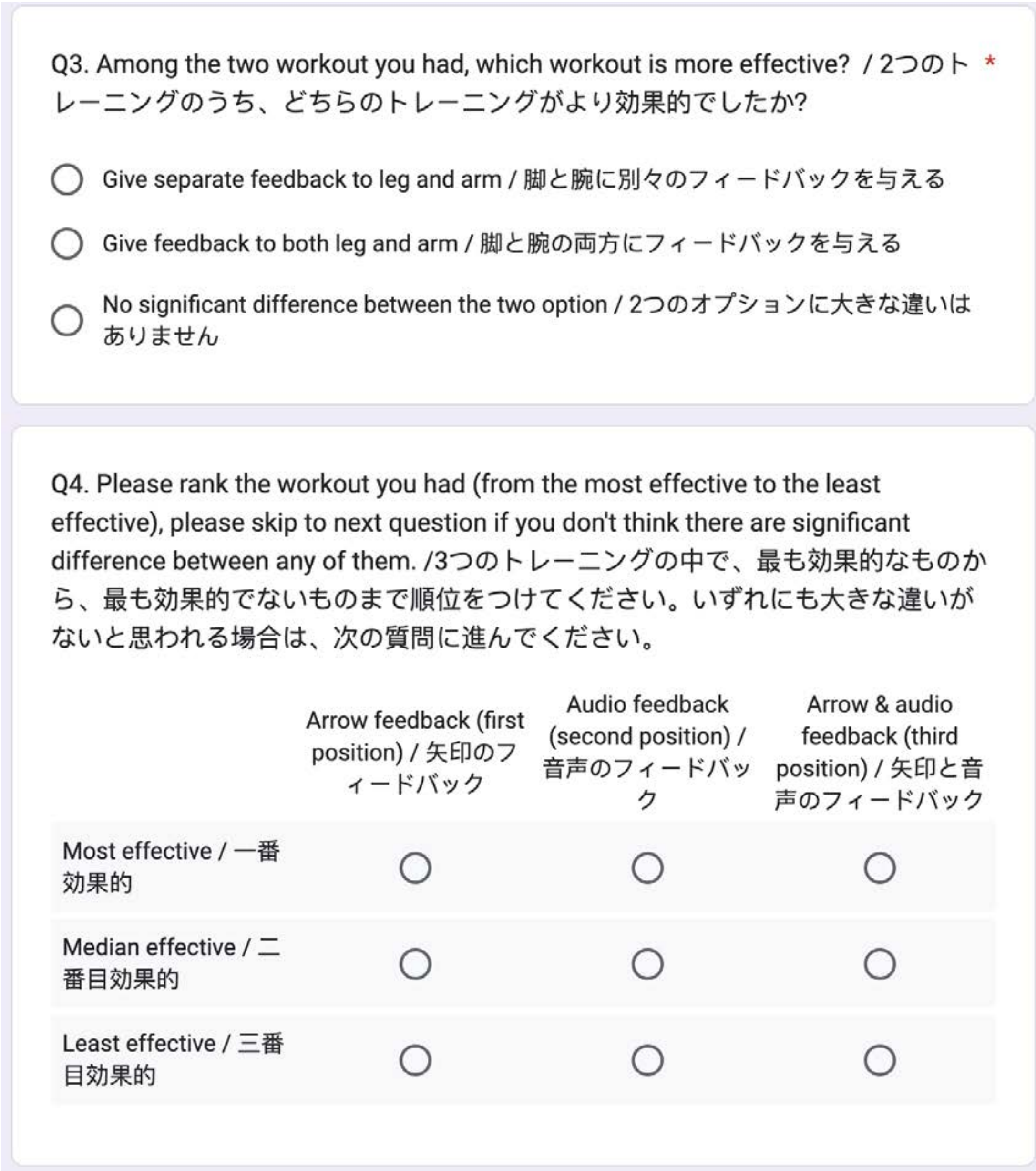
Disagree Strongly　　Neutral/ Mixed　　Agree Strongly

Q1. Compared to my normal workout routine that provides no feedback, I felt my position is closer to trainer's position during the workout / フィードバックがないトレーニングに比べ、トレーニング中の自分の姿勢はトレーナーの姿勢に近くなった気がする *

	1	2	3	4	5	6	7	
Strongly Disagree / 非常にそう思わない	○	○	○	○	○	○	○	Strongly Agree / 非常にそう思う

Q2. Compared to my normal workout routine that provides no feedback, I felt I spent less time to correct my position. / フィードバックがないトレーニングに比べ、姿勢を修正に費やす時間が短くなった気がする *

	1	2	3	4	5	6	7	
Strongly Disagree / 非常にそう思わない	○	○	○	○	○	○	○	Strongly Agree / 非常にそう思う



Q3. Among the two workout you had, which workout is more effective? / 2つのトレーニングのうち、どちらのトレーニングがより効果的でしたか? *

- ○ Give separate feedback to leg and arm / 脚と腕に別々のフィードバックを与える
- ○ Give feedback to both leg and arm / 脚と腕の両方にフィードバックを与える
- ○ No significant difference between the two option / 2つのオプションに大きな違いはありません

Q4. Please rank the workout you had (from the most effective to the least effective), please skip to next question if you don't think there are significant difference between any of them. /3つのトレーニングの中で、最も効果的なものから、最も効果的でないものまで順位をつけてください。いずれにも大きな違いがないと思われる場合は、次の質問に進んでください。

	Arrow feedback (first position) / 矢印のフィードバック	Audio feedback (second position) / 音声のフィードバック	Arrow & audio feedback (third position) / 矢印と音声のフィードバック
Most effective / 一番効果的	○	○	○
Median effective / 二番目効果的	○	○	○
Least effective / 三番目効果的	○	○	○

Q5. If you don't think there are significant difference between the three of them, please write down which feedback and which feedback has similar effectiveness? / 3つのオプションに大きな違いがないと思われる場合は、どのオプションとどのオプションが同様の効果を持つか、記入してください。

Your answer

Q6. Besides the audio and video feedback, what other feedback do you want to get from workout / 音声と矢印のフィードバック以外、欲しいフィードバックはありますか?

Your answer

Q7. Please write down any feedback toward the experiment / 実験に対するご意見・ご感想をお聞かせください

Your answer

B. Code of Prototype

B.1 Code of Uploading Pictures

```
def getPts(result,img_h,img_w):
    pts = []
    if result.pose_landmarks:
        for id, lm in enumerate(result.pose_landmarks.landmark):
            if id not in [15, 16, 27, 28]:
                continue
            cx, cy = int(lm.x * img_w), int(lm.y * img_h)
            pts.append([cx, cy])
    return pts

def judgePts(std_pts,cur_pts):
    if len(std_pts) == len(cur_pts):
        loss1 = np.sqrt(np.power(std_pts[0][0]-cur_pts[0][0],2) +np.power(std_pts[0][1]-cur_pts[0][1],2))
        loss2 = np.sqrt(np.power(std_pts[1][0]-cur_pts[1][0],2) +np.power(std_pts[1][1]-cur_pts[1][1],2))
        loss3 = np.sqrt(np.power(std_pts[2][0]-cur_pts[2][0],2) +np.power(std_pts[2][1]-cur_pts[2][1],2))
        loss4 = np.sqrt(np.power(std_pts[3][0]-cur_pts[3][0],2) +np.power(std_pts[3][1]-cur_pts[3][1],2))
        if loss1<thresh and loss2<thresh and loss3<thresh and loss4<thresh:
            return True
        else:
            return False
    else:
        return False
```

```
class Thread(QtCore.QThread):
    signal_logs = QtCore.pyqtSignal(bool)
    def __init__(self,parent=None):
        super(Thread, self).__init__(parent)
        self.ui = parent
        self.w = self.ui.label_video.size().width()
        self.h = self.ui.label_video.size().height()
        self.working = True
        self.stop_id = 0

    def stop(self):
        if self.working:
            self.working = False
            # self.wait()

    def restart(self):
        if not self.working:
            self.working = True
```

```
def run(self):
    # while self.working:
    cap = cv2.VideoCapture(self.ui.video_path)
    frame_count = int(cap.get(cv2.CAP_PROP_FRAME_COUNT))
    if cap.isOpened():
        for i in range(frame_count):
            if i<self.stop_id:
                continue
            if self.working==False:
                self.stop_id = i
            else:
                success,img = cap.read()
                if not success:
                    continue
                img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
                result = pose.process(img_rgb)
                if result.pose_landmarks:
                    mp_draw.draw_landmarks(img, result.pose_landmarks, mp_pose.POSE_CONNECTIONS)

                img_h, img_w, c = img.shape
                cur_pts = getPts(result, img_h, img_w)

                flag = judgePts(self.ui.std_pts, cur_pts)
                if flag:
                    # self.ui.lineEdit.setText("Correct")
                    cv2.putText(img, "OK", (10, 50), 1, 3, (0, 0, 255), 3)
                    cv2.imwrite("./temp/video.jpg", img)
                else:
                    # self.ui.lineEdit.setText("Wrong")
                    cv2.putText(img, "ERROR", (10, 50), 1, 3, (0, 0, 255), 3)
                    cv2.imwrite("./temp/video.jpg", img)

                self.ui.label_video.setPixmap(QtGui.QPixmap("./temp/video.jpg").scaled(self.w, self.h))
        self.stop_id = 0
    else:
        self.ui.label_video.setText("Failure on opening video")
```

B.2 Code of Build QMainWindow and Qdialog

```python
class Ui_Form(QtWidgets.QWidget):
    def __init__(self):
        super(Ui_Form, self).__init__()
        self.setupUi(self)
        self.connectInit()
        self.img_path = ""
        self.video_path = ""
        self.thread = Thread(self)
        self.std_pts = []

    def setupUi(self, Form):
        Form.setObjectName("Form")
        Form.resize(1030, 531)
        self.label_video = QtWidgets.QLabel(Form)
        self.label_video.setGeometry(QtCore.QRect(20, 20, 640, 480))
        self.label_video.setFrameShape(QtWidgets.QFrame.Box)
        self.label_video.setAlignment(QtCore.Qt.AlignCenter)
        self.label_video.setObjectName("label_video")
        self.label_img = QtWidgets.QLabel(Form)
        self.label_img.setGeometry(QtCore.QRect(680, 20, 331, 271))
        self.label_img.setFrameShape(QtWidgets.QFrame.Box)
        self.label_img.setAlignment(QtCore.Qt.AlignCenter)
        self.label_img.setObjectName("label_img")
        self.lineEdit = QtWidgets.QLineEdit(Form)
        self.lineEdit.setGeometry(QtCore.QRect(680, 310, 331, 31))
        font = QtGui.QFont()
        font.setPointSize(14)
        self.lineEdit.setFont(font)
        self.lineEdit.setAlignment(QtCore.Qt.AlignCenter)
        self.lineEdit.setObjectName("lineEdit")
        self.widget = QtWidgets.QWidget(Form)
        self.widget.setGeometry(QtCore.QRect(680, 420, 331, 31))
        self.widget.setObjectName("widget")
        self.horizontalLayout = QtWidgets.QHBoxLayout(self.widget)
        self.horizontalLayout.setContentsMargins(0, 0, 0, 0)
        self.horizontalLayout.setObjectName("horizontalLayout")
        self.pushButton_video = QtWidgets.QPushButton(self.widget)
        font = QtGui.QFont()
        font.setPointSize(14)
        self.pushButton_video.setFont(font)
        self.pushButton_video.setObjectName("pushButton_video")
        self.horizontalLayout.addWidget(self.pushButton_video)
        self.pushButton_img = QtWidgets.QPushButton(self.widget)
```

```
        font = QtGui.QFont()
        font.setPointSize(14)
        self.pushButton_video.setFont(font)
        self.pushButton_video.setObjectName("pushButton_video")
        self.horizontalLayout.addWidget(self.pushButton_video)
        self.pushButton_img = QtWidgets.QPushButton(self.widget)
        font = QtGui.QFont()
        font.setPointSize(14)
        self.pushButton_img.setFont(font)
        self.pushButton_img.setObjectName("pushButton_img")
        self.horizontalLayout.addWidget(self.pushButton_img)
        self.widget1 = QtWidgets.QWidget(Form)
        self.widget1.setGeometry(QtCore.QRect(680, 470, 331, 31))
        self.widget1.setObjectName("widget1")
        self.horizontalLayout_2 = QtWidgets.QHBoxLayout(self.widget1)
        self.horizontalLayout_2.setContentsMargins(0, 0, 0, 0)
        self.horizontalLayout_2.setObjectName("horizontalLayout_2")
        self.pushButton_start = QtWidgets.QPushButton(self.widget1)
        font = QtGui.QFont()
        font.setPointSize(14)
        self.pushButton_start.setFont(font)
        self.pushButton_start.setObjectName("pushButton_start")
        self.horizontalLayout_2.addWidget(self.pushButton_start)
        self.pushButton_stop = QtWidgets.QPushButton(self.widget1)
        font = QtGui.QFont()
        font.setPointSize(14)
        self.pushButton_stop.setFont(font)
        self.pushButton_stop.setObjectName("pushButton_stop")
        self.horizontalLayout_2.addWidget(self.pushButton_stop)

        self.retranslateUi(Form)
        QtCore.QMetaObject.connectSlotsByName(Form)
```

```
    def retranslateUi(self, Form):
        _translate = QtCore.QCoreApplication.translate
        Form.setWindowTitle(_translate("Form", "Form"))
        self.label_video.setText(_translate("Form", "Video(User)"))
        self.label_img.setText(_translate("Form", "Standard Action(Trainer)"))
        self.pushButton_video.setText(_translate("Form", "Upload Video"))
        self.pushButton_img.setText(_translate("Form", "Upload Action Picture"))
        self.pushButton_start.setText(_translate("Form", "Start"))
        self.pushButton_stop.setText(_translate("Form", "Pause"))


    def connectInit(self):
        self.pushButton_img.clicked.connect(self.slot_img)
        self.pushButton_video.clicked.connect(self.slot_video)
        self.pushButton_stop.clicked.connect(self.slot_stop)
        self.pushButton_start.clicked.connect(self.slot_start)
```

B.3 Code of Putting video and Image inside system by using RGB Camera and OpenPose

```
def slot_video(self):
    # fileName, _ = QtWidgets.QFileDialog.getOpenFileName(self, "select image", os.getcwd(), "All Files(*);;")
    fileName, _ = QtWidgets.QFileDialog.getOpenFileName(self, "select image", os.getcwd(), "(*.mp4)")
    if fileName != "":
        self.video_path = fileName
        cap = cv2.VideoCapture(fileName)
        if cap.isOpened():
            ret,img = cap.read()
            cv2.imwrite("./temp/video.jpg", img)
            w = self.label_video.size().width()
            h = self.label_video.size().height()
            self.label_video.setPixmap(QtGui.QPixmap("./temp/video.jpg").scaled(w, h))
            self.thread.working = False
            self.thread.stop()
        else:
            pass


def slot_img(self):
    # fileName, _ = QtWidgets.QFileDialog.getOpenFileName(self, "select image", os.getcwd(), "All Files(*);;")
    fileName, _ = QtWidgets.QFileDialog.getOpenFileName(self, "select image", os.getcwd(), "Images (*.jpg *.png)")
    if fileName != "":
        self.img_path = fileName
        img = cv2.imread(fileName)
        img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
        result = pose.process(img_rgb)
        img_h, img_w, c = img.shape
        self.std_pts = getPts(result, img_h, img_w)
        w = self.label_img.size().width()
        h = self.label_img.size().height()
        self.label_img.setPixmap(QtGui.QPixmap(fileName).scaled(w, h))
        self.lineEdit.setText("Standard Action(Trainer)")


def slot_start(self):
    self.thread.working = True
    self.thread.start()

def slot_stop(self):
    self.thread.stop()
```

B.4 Code of Comparing

```
if __name__=="__main__":
    mp_pose = mp.solutions.pose
    pose = mp_pose.Pose()
    mp_draw = mp.solutions.drawing_utils

    thresh = 30 #more strict the position is, the smaller the number will be

    if not os.path.exists("./temp"):
        os.mkdir("./temp")

    app = QtWidgets.QApplication(sys.argv)
    ui = Ui_Form()
    ui.show()
    sys.exit(app.exec_())
```