16th Mitsubishi Electric Cup Automation Contest

(Proposal for the Creative Design Category)

Name of entry:	“Light and Shadow” — An Intelligent Live Streaming Tracking Camera System
Team code:	vobtn7s5
Participating school:	Zhejiang University
Supervisors:	吴敏、張建良
Team members:	郭涛、劉芸雯、林俊如

July 2023

1. Project Background

According to the relevant reports, the amount of added value from the digital economy in China already exceeds 40 trillion yuan. The digital economy's contribution to the GDP is close to 70%, and is becoming a new driving force promoting China's economic development. Furthermore, live streaming is an important part of the digital economy. As of December 2022, there were more than 960 million network video users (including short video) in China, accounting for more than 94% of all internet users there.

The innovations of the live streaming economy as represented by "Live Broadcast +" support the overall development of the real and digital economies. They have incredibly great social value and innovative significance, and play an important role in the development of new industries, new categories of business, and new business models.

However, there are numerous problems with existing live streaming modes. The main problems faced by small teams with a smaller technological capacity but high market share are as follows. Without expensive and stable controllers and post-processing technology, realizing smooth and dynamic tracking shots that follow the movements of the object being live streamed is difficult. The video screen will not be stable, and capturing crucial movements and decisive moments is impossible. It is also difficult to suitably control rays of light during live streaming. Problems arise such as the screen blurring when the environment is too dark and not being able to see details, while overexposure is the result when the environment is too bright. All of this impacts the viewing experience. Furthermore, because the existing live streaming modes are drab and rigidly uniform, they are not able to completely realize the live streaming host's personalized shooting processes and their needs for unique modes of expression. They are unable to highlight the creativity of their unique personalities and labels among the countless live streaming teams.

To address these problems, our project makes the most high-precision and high-performance technological advantages, primarily those of Mitsubishi Electric's Q series PLC and host computers. Through intelligent tracking, intelligent fill lighting, and intelligent interactive technology, our system produces dynamic shooting screens by accurately recognizing the position and movements of the subject and automatically adjusting the camera's angle and focus while simultaneously making it possible to track line of sight. Also, through real-time analysis of ambient light, it makes intelligent adjustments to the brightness and angle of the light source and offers optimal conditions for the photographic subject. Furthermore, using intelligent interactive technology, it offers the user easy and intuitive operational procedures. At the same time, it also provides a completely new and upgraded live streaming shooting experience that supports the user with designing unique shooting modes and maximizes the limitless possibilities to pursue their own creativity.

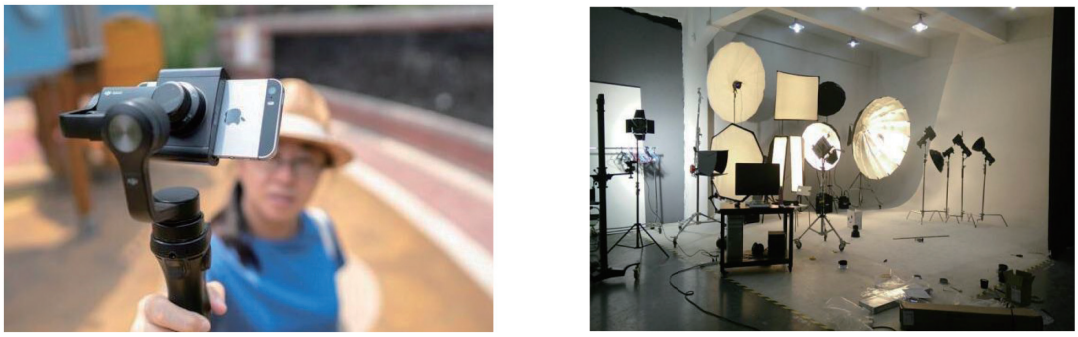

Fig. 1.1 Current state of the live streaming industry

2. Project Introduction

2.1 Project intent and overall objectives

This project is based on Mitsubishi Electric's automation products. We have made the most of Mitsubishi PLCs' high operational efficiency and abundant functional modules and designed a human-machine interactive system based on Mitsubishi touchscreens. We have also combined a range of technologies, including photographic subject detection, Ethernet transmission, and single

neuron PID control. Our objectives are to provide a completely new live streaming shooting experience for both live streaming teams and individual broadcasters. It will solve the problems they face, such as tracking the position and line of sight of the subject as well as offer effective control of light rays when streaming. Also, the unique live streaming presentation format designs and implements a creative model for personalized live streaming that will attract viewer interest. At the same time, it leaves room for rich software cultivation that will allow for more professional and intelligent shooting modes to be developed to handle user-designated scenarios.

2.2 The original aspects of the project

2.2.1 Line of sight tracking, screens with movement

The conventional shooting process requires the camera person to continually adjust the camera's position and angle to follow the subject. This would not only increase the complexity of the work but also constrain the degree of freedom in the shooting process. Conversely, this product, by making the most of the Mitsubishi PLC, precisely controls the many degrees of freedom of multiple motors and—thanks to its use of intelligent tracking technology—will accurately detect the positions and movements of the subjects photographed. Additionally, it controls the multiple degrees of freedom in the motor's horizontal plane and is able to track the subject in real time. Furthermore, to track the line of sight of the subject being photographed, it focuses on the subject's facial expressions and controls the motor's operations such as height and depth while automatically adjusting the camera's angle and depth of focus to track the subject's line of sight. To summarize, through the effective control of multiple motors through a PLC, the system can connect with dynamism each screen from the streaming shooting process. For all types of shooting, including live streaming, short videos, and other scenarios, the user can achieve smoother and more natural shooting effects.

2.2.2 Shines light, captures details

Rays of light are an important element that will have an impact on photographic effects. Depending on lighting conditions, problems may occur such as a darkened screen or overexposure. To solve these problems, this product uses intelligent fill-light technology. Based on a real-time analysis of the ambient light, it employs a single neuron PID control algorithm to calculate with OpenCV the brightness and color cast of the screen at hand while also finding the difference between the current brightness and the color temperature of the supplemental lighting. It then uses this latter difference as the control amount for the proportional, integral, and derivative stages of the single neuron PID control algorithm. Furthermore, we use Hebb learning algorithm to obtain the control weightings. We then use a digital-analog converter to take the final output values and convert them to measurements of the current that controls the supplemental lighting. In so doing, we track in real time the luminosity and color temperature of the shooting environment, and deliver the optimal lighting conditions for the subject.

2.2.3 Joining hearts, innovation without limits

Videography and live streaming calls for flexible operations and diverse features that will satisfy the user's creative needs. For this product, we have developed an intelligent interactive technology based on Mitsubishi touchscreens. With it, we offer easy and intuitive operating procedures that allow users to readily learn and employ a variety of features. The design of the still shooting mode allows the user to select their desired shooting angles. Also, the design of the dynamic shooting mode offers multiple standard modes for the user to select. For example, the shooting partner mode is suited to shooting in expansive outdoor settings because it keeps the angle of the camera pole uniformly constant and the device automatically tracks the user. The user can also customize and freely set routes of travel and fill-light modes to fully express their own creativity.

2.3 Consolidation of resources and system design

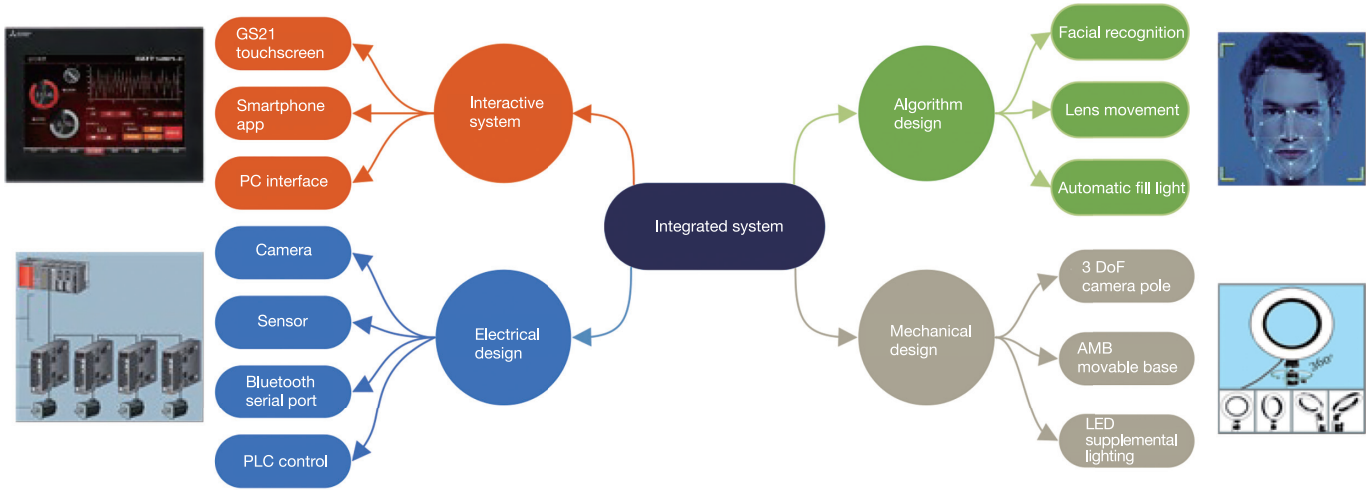

Fig. 2.1 Integrated system block diagram

This project, based on the Mitsubishi MELSEC-Q series PLC and host computers, combines such technologies as facial recognition, Ethernet transmissions, and PID control to deliver to the user a variety of solutions such as multi-angle shooting, automated tracking shooting, and automated supplemental lighting.

2.3.1 Accurately controls motor rotation, accurately adjusts shooting angles

The Mitsubishi MELSEC-Q series PLC features fast response speeds, superior control precision, and an abundant variety of ports. The built-in PWM module provides precise control of multiple rotating stepper motors and of supplemental lighting and other such devices. With all of this, it can reduce the time needed to adjust the angle of camera poles. The Q-series PLCs also support Ethernet connections. This helps to ensure a highly reliable TCP/IP connection with the host computer and AMB base and improves the system's reliability and accuracy.

2.3.2 Handles multiple complex environments with two types of wireless control solutions

Considering the complexity of the shooting and live streaming environments, this project provides users with two different operational methods. In one method, the user employs a UI interface on the host computer to exchange data with the PLC via Ethernet. It has the merit of being fast and reliable, and is well suited to indoor live streaming environments. With the other method, employing a Bluetooth serial port module the user interacts with the PLC via a smartphone app. This method has the merits of simple operation and being capable of real-time movement. It is well-suited to outdoor live streaming environments.

Fig. 2.2 Bluetooth-based app for connecting with the host computer UI interface and smartphones

2.3.3 Building a personalized live streaming solution with multiple features that users can select

Different types of live streaming will mean differences in shooting angles and shooting paths. Different users will mean different shooting needs. In this project, we offer multiple live streaming modes. With the still shooting mode, the user can select their desired shooting angle. With the dynamic shooting mode, we provide multiple standard modes among which the user can select. For example, the shooting partner mode is suited to shooting in expansive outdoor settings because it keeps the angle of the camera pole uniformly constant and the device automatically tracks the user. The user can also customize and freely set routes of travel and fill-light modes to fully express their own creativity.

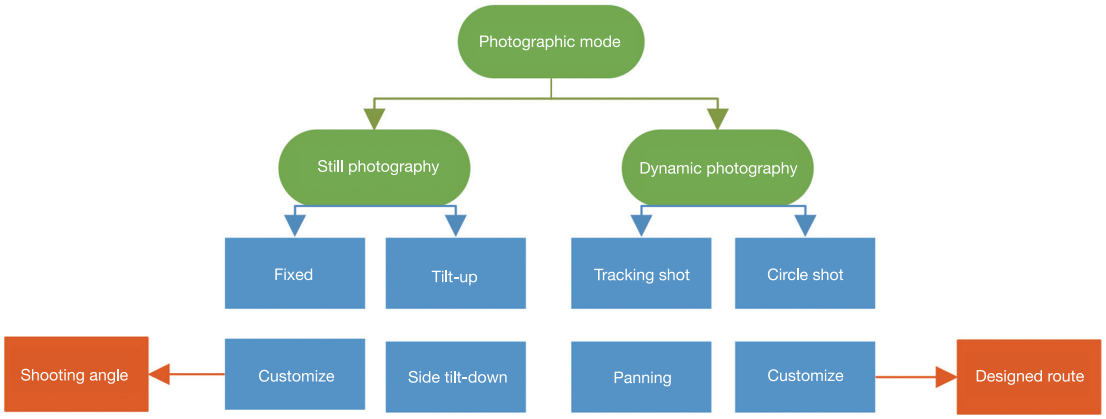

Fig. 2.3 Multiple photographic modes

2.3.4 Superlative mechanical structure design, stable and vivid shooting screen

This project uses a tapered construction with an AMB movable base with two-wheeled differential gearing. It offers an extremely high-tech look that addresses users' aesthetic needs in the live streaming industry, and maintains the stability of the captured image. With this, we have managed to shift the simple appearance and center of gravity of the entire device downward. Also, to help users secure their live streaming devices and improve their experience, we integrated a smartphone anchoring device and a supplemental lighting device.

2.4 Adoption of Mitsubishi Electric technology

2.4.1 PLC-based target tracking and effective control of light

The MELSEC-Q series PLC we used in this project features fast response speeds and superior control precision. It responds quickly to outside commands and can improve the shooting precision of the device as a whole. Also, the internally integrated PWM output and timer modules accurately control the rotational angles and speeds of the stepping motors. Furthermore, the PLC records the motor's rotation speed and protects the mechanical structure from damage due to impacts.

2.4.2 MC protocol realizes Ethernet transmissions among multiple devices

The Mitsubishi PLC support Ethernet transmissions. This MC transmission protocol helps to achieve information interaction between the PLC and the GS21 series touchscreen, host computer, and AMB movable base. The host computer program can bypass complex bottom layer transmission protocols, directly call the API interface, and read and write internal PLC data. It needs only network switches and network cables to realize this. For these reasons, compared to a serial port connection, it is easy to operate and more convenient, and follow-up maintenance is also easier.

2.4.3 Achieving excellent interactive control performance based on a Mitsubishi Electric touchscreen

The Mitsubishi GS21 series touchscreen supports screen design and data interaction with PLC transmissions is easy. Accordingly, it provides the user with a rich screen and also offers an easily controlled human-machine interactive interface. Employing the GT Designer3 software allows the user to quickly design easily manipulated operational pages that offer high screen resolution and little excess.

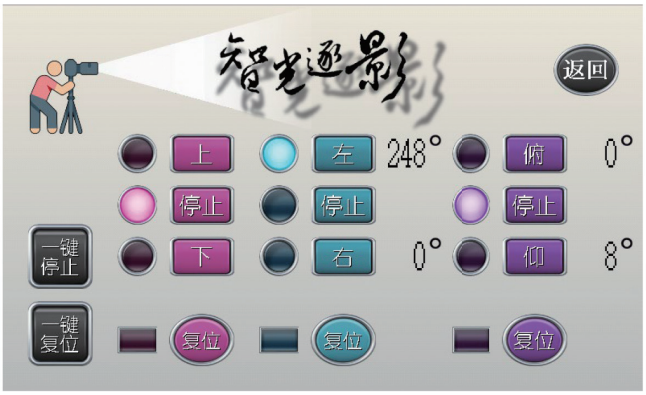

Fig. 2.4 Interactive interface for the GS21 series touchscreen

3. Application Analysis

3.1 Current state of market development

With the spread of a mobile internet and the promotion of 5G networks, the live streaming and short video markets are currently in a stage of rapid development. The live streaming economy is one that is based on the injection of human resources. It mainly takes the form of broad-ranging and collective human interaction, and the sharing of information and feelings make up its basic content. The earning of direct or indirect economic income based on the acquisition of interpersonal resources is an important part of the digital economy.

Since 2016, the live streaming economy has shown vigorous development in China. In 2017, the total GMV of live streaming transactions in China reached 31 billion yuan, and in 2019 it reached 390 billion yuan. Additionally, more than 40 live streaming companies had invested more than two billion yuan. Still further, as of June 2019 there were already some 433 million people using networked live streaming in China.

This rapid development of the market has produced both problems and opportunities. More and more companies and people are engaging in live streaming and creating short videos, leading to ever more intense competition. Furthermore, as user needs for quality and innovative content rise, superior content and originality are becoming important elements for drawing in users and fans. Accordingly, to respond to changing market needs, creators and platforms must continue to innovate and improve the user experience.

3.2 Potential market value

a. Realizing creators' dreams: Individual streamers, content creators, and social influencers need high quality photography tools to fulfill their dreams of creation and show off their unique personalities and talents.

b. Branding concerns: Companies and brands want to stand out in extremely competitive marketplaces. The promotion of products and services through live streaming and short videos is an industry concern.

c. Interactive education portal: The online education and training industry could open a new portal for education by providing vital interactive learning experiences to students via high-quality video content.

d. Don't overlook a decisive moment: Active live streaming and sports coverage can deliver supreme pleasure to viewers because it can capture decisive moments by means of intelligent shooting tools and broadcast them to viewers in real time.

3.3 Future forecast

a. Growing market needs: With the continued development of the live streaming and short video industries, the need for creative, high quality content is also expected to grow. "Light and Shadow" is outfitted with various features including intelligent tracking, intelligent fill-lighting, and intelligent interactiveness. We expect it will be able to fulfill the advanced shooting needs of creative and professional users, and win support from an even greater number of users.
b. Expanding the areas of application: "Light and Shadow" has the potential for use even in industries other than those of live streaming and short videos. For example, it is possible that people will need intelligent photographic devices to record and stream content in such fields as education, training, sports, and entertainment among others. It will also be possible to further expand the product's market share by entering new industries and areas.
c. User customization needs: As user needs continually change, the trend to greater personalization and customization is likewise growing. "Light and Shadow" imagines that, in order to satisfy specific user needs and be more competitive, it will deliver even more customization options and personalized features.
d. Combines with other industrial products under the Mitsubishi umbrella: By replacing motors with different degrees of freedom with more powerful Mitsubishi servo motors, this type of intelligent tracking photography product achieves more powerful movements and precise controllability. This product can be used not only for live streaming and short video creation, but also for the monitoring and capturing of production processes in the automobile industry. In production, to preserve production quality and efficiency, it is necessary to monitor and record numerous complex procedures. An intelligent tracking photography product can automatically track and capture important procedures in the vehicle assembly and inspection and measurement processes on the production line thanks to the high output and precise control capabilities of Mitsubishi servo motors. Such application scenarios will be beneficial in many ways to automobile manufacturers as shown below. For example, automated photography will greatly cut back on human resources and cut monitoring costs. By shooting crucial operations and details in the vehicle production process, the intelligent tracking feature will monitor product quality and quickly discover problems. Intelligent fill-light technology can ensure well-lit photographic scenes even in the manufacturing workplace where adequate amounts of light cannot be guaranteed.

4. Technical Feasibility Analysis

4.1 Mechanical design

The mechanical design of this project comprises a movable base, support platform, and a 3 DoF shooting platform. A 1:1 sample of the completed product is as follows.

Fig. 4.1 Completed sample

4.1.1 Support platform

Used to mount an electrical module, and provides reliable support for the 3 DoF shooting platform. Sample is as follows.

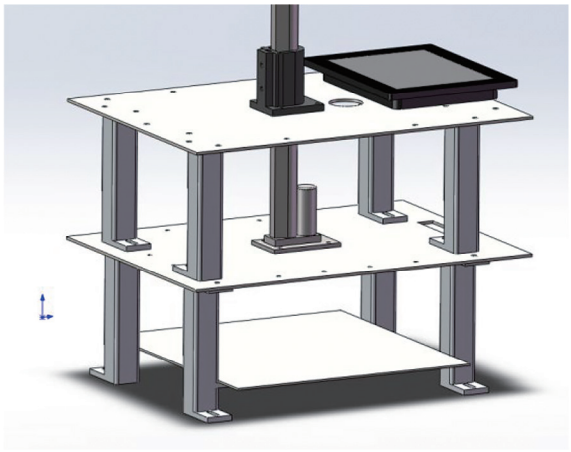

Fig. 4.2 Support platform

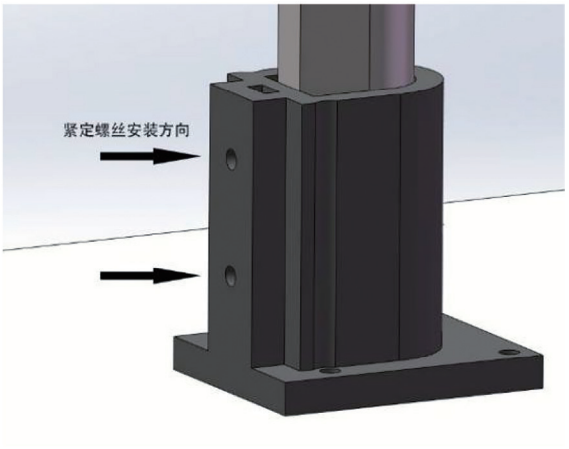

Fig. 4.3 Limit module

The support platform is split up into three levels. The bottom layer is used to mount crucial larger electrical devices such as inverters and stabilized switching DC power supplies. The middle layer is used to house various small electrical circuit devices such as the Mitsubishi PLC and stepping motor drive modules, supplemental lighting drive panels, and the like. It also has mounting holes for the elevating rods used for the 3 DoF shooting platform.

The top layer of the support platform is used for installing Mitsubishi's industrial touchscreens and PCs. Also, the central limit module offers further bearing capacity for the 3 DoF shooting platform's elevating rod. It prevents the photography platform from rocking as the camera moves with tracking.

As shown in Fig. 4.3, the limit module uses two fixing screws to position the elevating rod in the center. For that reason, it is able to alleviate to the greatest extent possible the rocking to the 3 DoF shooting platform generated when it being moved.

4.1.2 3 DoF shooting platform

The 3 DoF shooting platform is the core element of this product. A sample is shown on Fig. 4.4

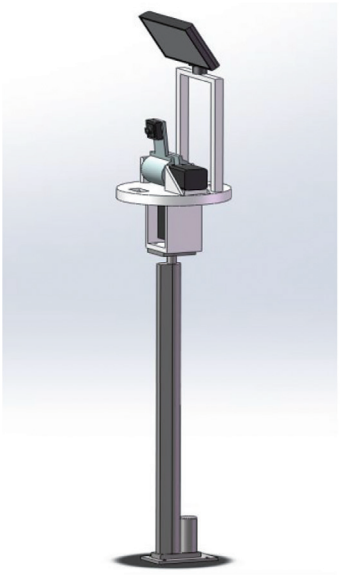

Fig. 4.4 3 DoF shooting platform

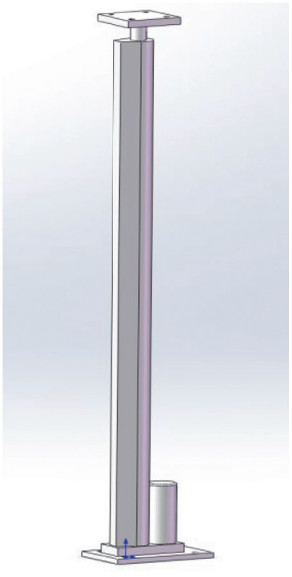

Fig. 4.5 Elevating rod

The 3 DoF shooting platform integrates multiple features. It can use smartphones from a variety of manufacturers and with various specifications as photographic devices. It offers three degrees of freedom when shooting—lift-up/lift-down, pan-roll, and tilt-up/tilt-down—and so basically covers all shooting settings. The angle of the supplemental lighting on top can be changed to provide different degrees of intensity and of color temperatures and make up for inadequate ambient light.

The degree of lift-up/lift-down freedom for the shooting platform is provided by a single elevating rod. An example of what this looks like is shown in Fig. 4.5. The DC motor for the elevating rod is decelerated using the gear set and its torque is amplified. Following this, by driving the rotation of a single screw rod, the upper platform of the elevating rod is able to perform the lift-up/lift-down function. The elevating rod offers a lift-up/lift-down speed of 5 cm/s and up to 20 kg of driving force, guaranteeing free lift-up/lift-down movements for the shooting platform.

The shooting platform's pan-roll and tilt-up/tilt-down degrees of freedom are accomplished by linking together two 42-mm stepping motors. A sample of this machine's structure is presented on Fig. 4.6.

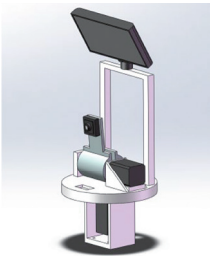

Fig. 4.6 Shooting platform (pan-roll and tilt-down/tilt-up section)

The entire structure for affixing the smartphone is cylindrical in shape. There is a square slot in the center that can accommodate different models of smartphone. The subject detection camera that is used for tracking shots is also attached here. This guarantees that the detection screen is extremely close to the screen actually shot by the smartphone.

The "doorframe-type" support for the shooting platform is used for attaching supplemental lighting. A connection design that mimics that of a joint is used to connect the supplemental lighting and the shooting platform to ensure that the direction of the supplemental lighting can be freely adjusted as needed.

4.2 Electrical design

The system achieves all of its functions with a MELSEC-Q series PLC as its core processor; a smartphone app, GS21 series touchscreen, and host computer as its human-machine interactive channel; a distortion-free wide-angle camera, laser sensors, and so forth that provide the forward channel; and a rotation motor, telescoping pole, AMB movable base wheel motor, and supplemental lighting that serve as the rearward actuator.

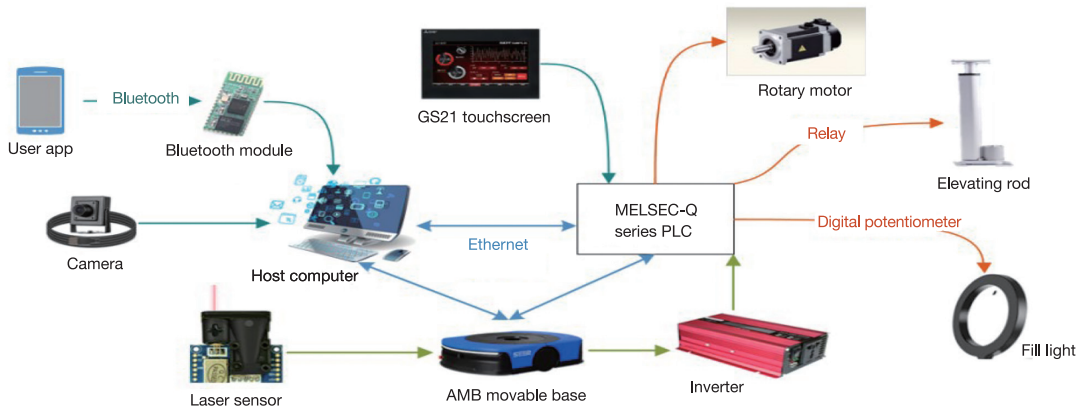

Fig. 4.5 Electrical design system diagram

4.3 Algorithm design

4.3.1 Machine vision—Face tracking

After startup, the user positions themself directly in front of the camera. After startup, the host computer calls the OpenCV and launches the face detection and tracking programs. First, it compresses the image and converts it to grayscale. It also loads the harr operator and finds the face within the frame screen. When the face and model are matched, the host computer acquires the keypoint coordinates for the face and transmits the corresponding drive signal information to the MELSEC-Q series PLC by Ethernet. If matching fails, the above procedures are repeated.

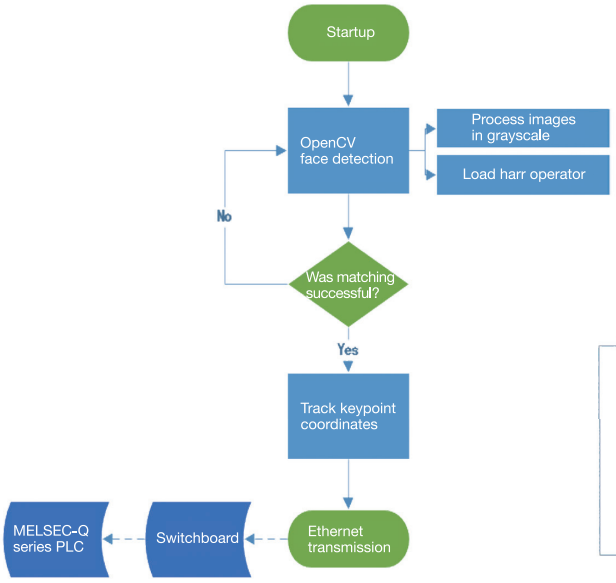

Fig. 4.6 Face tracking algorithm flowchart

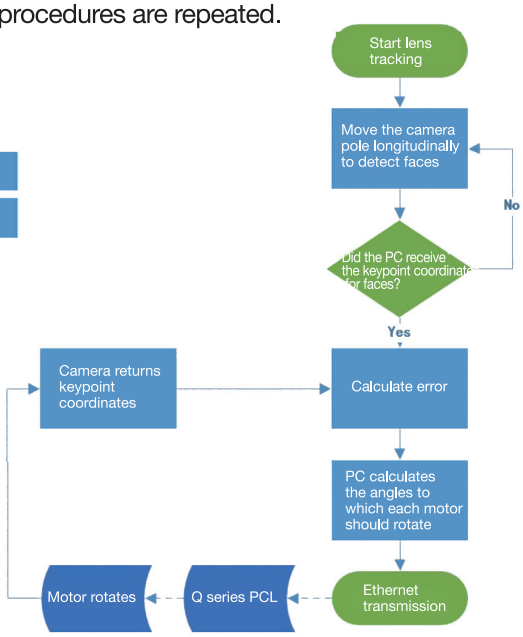

Fig. 4.7 Lens tracking algorithm flowchart

4.3.2 Lens tracking: PLC control

The control algorithm that allows the 3 DoF shooting pole to achieve lens tracking is as shown in Fig. 4.7. When the mode selected by the user activates the lens' tracking function, the user will need to face the camera directly. The host computer then calls up the Mediapipe library, detects the keypoint of the subject, and returns the keypoint coordinates. The distance between the target and device is predicted based on the principles of monocular ranging, and furthermore the deviation from the target value is calculated. The information about the angle to which the motor should rotate is transmitted by Ethernet to the PLC. By sending a constant pulse signal and controlling the motor so it rotates at the specified angle, the PLC ensures that the face is always positioned at the center of the screen.

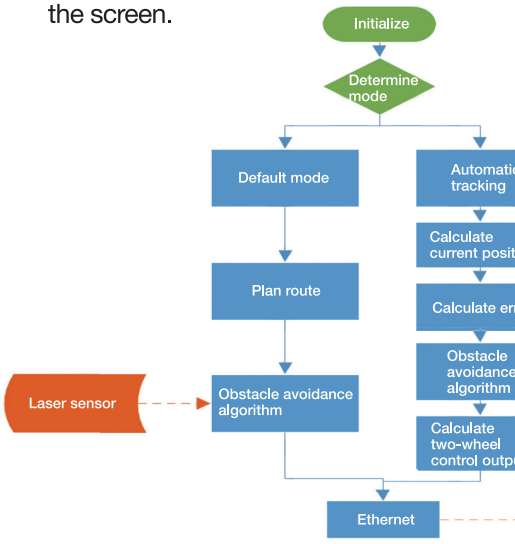

Fig. 4.8 Flowchart of base mobility algorithm

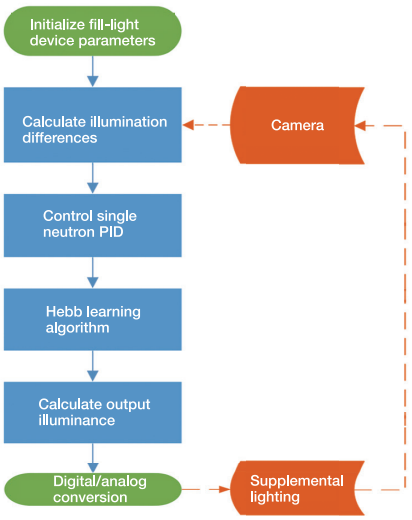

Fig. 4.9 Flowchart of the single neutron PID control algorithm

4.3.3 Movements of the base, closed-loop control

The control algorithm for the AMB movable base is as shown in Fig. 4.8. On device startup, first the parameters for the AMB movable base are output. The PLC then determines what mode the user has selected. If it is default mode, the PLC will combine the AMB's obstacle avoidance algorithm and the anti-shake algorithm to calculate the current position based on the screen captured by the camera and the Mediapipe library's subject detection. It then plans a travel route and controls the rotation speed and rotation direction of the wheel motor by Ethernet. If it is dynamic mode, it keeps the shooting pole stationary, calculates the current position based on monocular ranging and the keypoint coordinates returned from the camera, subtracts that from the target's position to determine the error, and adjusts the output of the wheel motor as a feedback quantity to realize the device's tracking functions.

4.3.4 Automatic fill light: Single neutron PID control algorithm

As shown in Flowchart 4.9, this device uses a single neuron PID control algorithm that makes it possible to instantly change the brightness of the supplemental lighting according to the brightness of the environment. First, it uses OpenCV to calculate the brightness and color cast of the current screen and determines the difference between the supplemental lighting and the current brightness and color temperature. This difference is used for the proportional, integral, and derivative stages of the single neuron PID control algorithm. A Hebb learning algorithm is further used to obtain the control weighting. The output value that is finally obtained is converted using a digital-analog converter to the measured value for the current amount that controls the supplemental lighting. Doing the foregoing realizes instantaneous tracking of the brightness and color temperature of the shooting environment.

5. Expected Results

1. Intelligent tracking and real-time fill lighting: A new upgrade to the shooting experience

This product's dynamic tracking feature captures smooth and crisp images as if they were photographs. It makes it possible users to freely change shooting angles on social media and live streaming platforms, allowing users to present vivid stories and instants. Its intelligent analysis and fill-light technology deliver bright and vivid photographic effects that will make the user's work more appealing. In the fiercely competitive live streaming and short video fields, this product provides creators who need visual effects to draw people's attention with superior features in order to light up the world with their creations.

2. The whereabouts of inspiration, the pursuit of quality:The end of limits on the creative journey, opening up a new era for quality photography

This product's ease of operation and innovative features make it possible for creators to open the way forward down which inspiration of their creations will go, and allow them to explore limitless possibilities on the sea of creation. In the live streaming and short video fields, users need tools that will satisfy their powers of creation. Thanks to its ease of operation and innovative features, this product will make it possible for users to more freely explore and express their creativity. The high-quality photographic effects and innovative features will produce memories that are difficult to forget. In the fiercely competitive content marketplace, user needs with respect to high-quality live streaming and short videos are growing with each passing day. This product's amazing photographic effects and innovative features will deliver a high-quality experience and leave precious memories in users' hearts.

3. Laying out creativity for the future, getting everyone to create: Building a shooting template shared platform, creating a creative trend made by everyone

In the future, we plan in conjunction with this product to build a universal shooting mode development platform for everyone and make templates of innovative shooting modes. Through an app, users will be able to develop their self-designed shooting templates and upload them to the development platform for sharing with other users. In the event that other users want to use the same shooting mode, they will be able to obtain the same mode with a single click. The platform will make it possible to offer a shooting guide that corresponds to shooting modes so the user can more easily obtain the same photographic effects. Through this platform built for users, users will be able to take full-blown videos without having any specialized basic knowledge and train themselves to be their own private, exclusive cameraman. With such a model that promotes the creative work of everyone, it can be expected that a further new dimension of trends in creative work will inevitably emerge.

[1] 丁梦寒,郑腾.温室大棚 LED 智能调光的补光装置设计[J].江苏农业科学,2021,49(23):201-206.
[2] 张慧,丁锦宏.基于 PLC 的四轮驱动室内移动机器人设计[J].江苏工程职业技术学院学报,2019,19(02):19-22.
[3] 乔敬雯.新时代直播电商行业的发展前景研究[J].中国商论, 2023, 879(08): 47-51.
[4] 马书雅.短视频的传播优势及其对摄影行业的影响[J].西部广播电视, 2022, 43(04): 63-65+81.
[5] 三菱 Q 系列可编程控制器使用说明书.