Spotlight Tracking with Automated Recognition (S.T.A.R.)

Gage Anderson, Jerison Lau, Travis Nguyen, Lucio Ruben Villena

CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, Florida 32816-2450

Dept. of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida, 32816-2450

The Spotlight Tracking with Automated Recognition (S.T.A.R.) system is an autonomous stage lighting fixture that tracks performers without a dedicated operator or lighting console. A CMOS camera and machine learning algorithm detect and follow subjects in real time, while a two-axis gimbal directs a focused beam. The illumination system is based on an ellipsoidal reflector spotlight architecture, using a reflector, gate, and BK7 lens to produce a controlled beam at ~7.5 m. The system integrates embedded processing and wireless control, demonstrating a compact, cost-effective solution for automated theatrical lighting.

***Index Terms* — CMOS image sensors, Computer Vision, High intensity discharge, LED lamps, Object tracking, Retroreflectors, Servomotors**

I. Introduction

Stage lighting plays a critical role in directing audience attention and enhancing visual storytelling in theatrical and live event environments. Traditional spotlighting methods rely either on trained operators manually controlling large fixtures or on complex lighting consoles that require significant cost and expertise. While these approaches are effective, they introduce limitations in consistency, accessibility, and scalability, particularly for smaller productions or educational settings.

Recent advances in computer vision and embedded systems have enabled new possibilities for automation in optical and imaging-based applications. Systems capable of detecting and tracking human motion in real time are now widely used in surveillance, robotics, and interactive media. Applying these technologies to stage lighting presents an opportunity to simplify operation while maintaining professional-level performance.

Motivation – The goal of the S.T.A.R. project is to develop a low-cost, autonomous spotlight system that can reliably track a performer in real time. Existing tracking systems often rely on RF-based tags or external camera networks, which can be expensive, complex to install, and prone to interference in live production environments. By using a single onboard imaging system and computer vision algorithms, S.T.A.R. aims to provide a simpler and more accessible solution.

Problem Statement – Current spotlight tracking solutions are either labor-intensive, requiring dedicated operators, or technologically complex and costly, limiting their use in smaller venues. Many systems also lack flexibility or require additional infrastructure such as multiple cameras or wearable tracking devices. A compact, self-contained system capable of real-time human tracking and automated beam control is needed to reduce cost, complexity, and setup time.

Objectives – The S.T.A.R. system is designed to detect and track human subjects using a CMOS imaging system and machine learning algorithms, and to translate this information into real-time control of a motorized gimbal for spotlight positioning. The system integrates optical, electrical, and software subsystems to produce a stable, high-intensity beam with consistent performance. The overall objective is to demonstrate a standalone, automated lighting system that is efficient, reliable, and suitable for theatrical and live event applications.

II. Overview of System Design

The S.T.A.R. system operates by combining real-time imaging, motion control, and optical projection to autonomously track and illuminate a performer. The system is designed as a standalone unit that does not require external tracking infrastructure or a lighting console. Instead, it integrates all major subsystems into a compact platform that performs detection, tracking, and beam steering in a continuous control loop.

The overall design is composed of four primary subsystems: the illumination system, imaging system, mechanical and motor control system, and embedded processing and software control. Each subsystem is developed to operate independently while maintaining tight synchronization with the others to ensure accurate tracking and stable optical output.

A. Illumination System Architecture

The illumination system is based on an ellipsoidal reflector spotlight (ERS) design, which efficiently collects and shapes light from a compact source. A high-power LED COB module serves as the illumination source, emitting light in every direction. This light is collected by a cold mirror ellipsoidal reflector, which directs visible wavelengths toward the optical axis while allowing infrared radiation to pass through the rear, reducing thermal load on downstream components.

The reflector forms an intermediate image at the gate, where beam shaping elements such as gobos or apertures can be placed. A BK7 biconvex focusing lens is positioned after the gate to project the beam onto the stage with a controlled output angle of approximately 19°. The system is designed to produce a uniform, well-defined spot at a distance of approximately 7.5 meters while minimizing optical aberrations and stray light.

B. Imaging System

The imaging system is responsible for detecting and tracking the performer within the scene. It consists of a CMOS camera module paired with a multi-element lens system designed to capture a wide field of view while maintaining sufficient resolution for accurate target detection.

Unlike the illumination system, the imaging system remains stationary and observes the entire stage continuously. This design allows the system to detect true motion of the subject without artificially re-centering the image. The lens system uses achromatic doublets to reduce chromatic and spherical aberrations, ensuring consistent image quality across the field.

The captured image is processed in real time to extract the position of the selected subject, which is then used to drive the motor control system.

C. Mechanical and Motor Control System

The mechanical system consists of a two-axis gimbal that controls the direction of the spotlight. It provides pan and tilt motion to follow the performer across the stage.

Servo motors are used for smooth and precise movement, with PWM control signals generated by a motor controller. The structure is designed to support the optical system while remaining lightweight and stable during operation.

D. Processing and Control System

SOFTWARE ARCHITECTURE

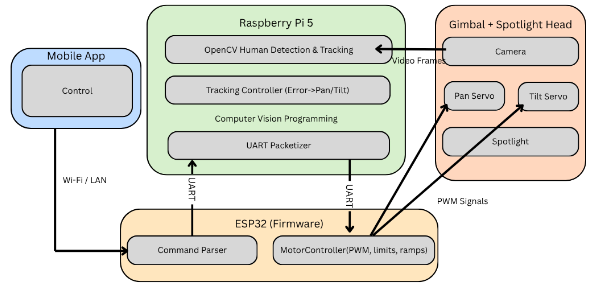

Fig 1. Software architecture of the S.T.A.R. system, illustrating computer vision processing on the Raspberry Pi, command transmission to the ESP32, and real-time control of the gimbal-mounted spotlight

The processing and control system coordinates performer detection, wireless user interaction, and real-time motor actuation. It is built around a distributed embedded architecture consisting of a Raspberry Pi Compute Module 5 (CM5), an ESP32-S3 microcontroller, and a mobile application interface. The Raspberry Pi performs the computationally intensive vision tasks, including image capture, person detection, and target localization, while the ESP32 handles deterministic control tasks such as mode management, command arbitration, and servo actuation. This division allows the system to maintain responsive tracking behavior while separating high-level perception from time-sensitive embedded control.

Communication between the processing nodes is designed for low latency and predictable timing. Tracking data generated by the Raspberry Pi is transmitted to the ESP32 over a UART link, where it is converted into pan and tilt commands for the motor control subsystem. The mobile application connects to the control system through a wireless interface and provides user access to operating functions such as power control, mode selection, target selection, and manual override.

III. SYSTEM COMPONENTS

S.T.A.R. operates using integrated optics, electronics, and computational subsystems. Each component provides a key role in the overall system's operation in a compact form. The key components and their role in S.T.A.R. are explained below.

A. Light Source

The illumination system of S.T.A.R. uses the ETC Source 4WRD II LED light engine as its primary light source. It is designed as a replacement for tungsten-halogen lamps in ellipsoidal reflector spotlight (ERS) systems, providing similar beam characteristics with improved efficiency and lower thermal output. The module produces high-intensity white light with a color temperature of approximately 3200K, an optimal choice for theatrical applications [2].

The light engine is placed at the first focal point of the ellipsoidal reflector to allow efficient light collection and projection through the optical system. Compared to incandescent sources, it operates at lower temperatures and has a longer lifetime, reducing thermal stress and maintenance while maintaining consistent optical performance.

B. Illumination Optics

The illumination optics of S.T.A.R. are based on an ellipsoidal reflector spotlight (ERS) design, consisting of a cold mirror reflector, gate, and focusing lens. Light from the source is placed at the first focal point of the reflector, which redirects and concentrates the visible light toward the gate while allowing infrared radiation to pass through for thermal management.

At the gate, the system forms an intermediate image where beam shaping elements such as apertures or slides can be inserted. A BK7 biconvex lens positioned after the gate projects the beam onto the stage with a controlled output angle of approximately 19°. This configuration enables efficient light collection, uniform beam distribution, and minimal optical aberrations.

Zemax Illumination System Diagram

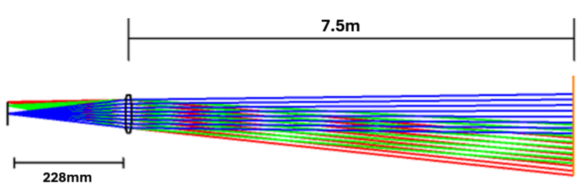

Fig 2. Ray-trace diagram of the illumination system output. The leftmost plane represents the aperture that is placed at the focal point of the reflector.

C. Raspberry Pi Compute Module 5

The Raspberry Pi CM5 serves as the vision-processing node of the S.T.A.R. system. It receives image data from the onboard camera and executes the computer vision pipeline used for person detection and target tracking. For each frame, the software identifies the selected performer, determines the location of the target within the field of view, and generates compact tracking telemetry representing position and tracking state. This information is then transmitted to the ESP32 for real-time actuation. By dedicating the Raspberry Pi to compute-intensive image processing, the system can support machine learning-based tracking without burdening the real-time motor-control layer. The imaging system along with the major system components are shown in Fig. X

Imaging System Optical Schematic

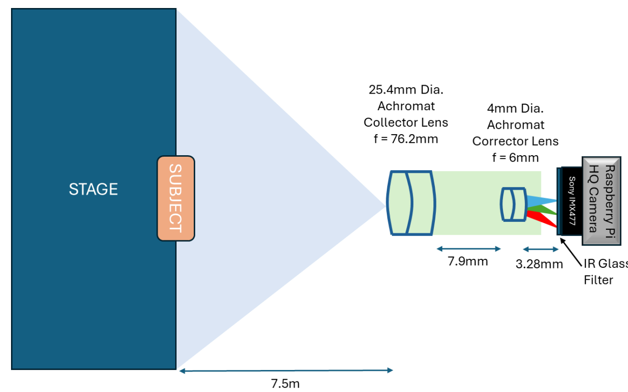

Fig. 3. Optical schematic used to project the entirety of the stage onto the CMOS sensor with labelled components.

D. CMOS Camera Module

The imaging system of S.T.A.R. uses a Raspberry Pi High Quality (HQ) Camera with a 12.3-megapixel Sony IMX477 sensor with a pixel size of 1.55µm × 1.55µm. With a CMOS sensor, the system does not require cooling or high input power, making it ideal for the simplicity of S.T.A.R. The sensor uses a rolling shutter to meet a low cost and compact size while also capturing high-dynamic range for stronger low-light performance.

E. Lens Design and Assembly

A system to achieve a 1250x demagnification from ~7.5m away consisted of two achromatic doublets spaced in air. Through iterations of the design, a two lens system as opposed to a three lens system was chosen to reduce costs of the overall system despite using commercial off the shelf components. The lens system has specialized coatings to reflect the wavelengths beyond and below the visible wavelength regime permitting only the wavelengths of 400nm to 700nm. The first lens facing the stage is the achromat collector lens with a 25.4mm diameter and a 76.2mm focal length. This lens provides the collection of the light from the target of 7.5m away with the accurate optical quality desired. The second lens of choice is a 4mm diameter with a 6mm focal length to provide the correct system focal length and magnification to the CMOS sensor.

ZEMAX IMAGING SYSTEM DIAGRAM

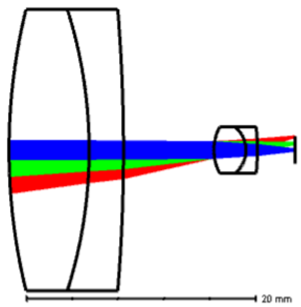

Fig. 4. Ray-trace diagram of the imaging system with the objective being 7.5m on the left of the first lens and the CMOS on the right-hand side.

Fig 4. demonstrates the overall system assembly. With an objective 7.5m away, it is not featured in the diagram to provide more emphasis to the distance between the lenses and the sensor. The light from the stage alongside the subject projects through the subsequent lenses and provides an image onto the Sony IMX477 sensor.

F. Custom ESP32-S3 Board

A custom-designed ESP32 MCU board based on the ESP32-S3-WROOM was developed for the S.T.A.R. system. The ESP32-S3 functions as the embedded control core of the system. It receives user commands from the mobile application, interprets tracking telemetry from the Raspberry Pi, and generates the final control outputs that drive the pan and tilt servos. In addition to motor actuation, the ESP32 manages system state transitions between idle, autonomous tracking, and manual control. This centralized state authority ensures that user commands, tracking updates, and safety responses are handled consistently. If telemetry becomes stale or the target is lost, the ESP32 can enforce a safe fallback behavior such as holding position, returning to a neutral state, or waiting for reacquisition.

G. Custom PCA9685 PWM Board

A custom PCB was developed around the PCA9685 IC. This board provides the necessary PWM control our servo motors need to track the human subject. Without this board, our MCU would not be able to communicate with the motors in our system, rendering them completely useless.

H. 600W AC/DC Power Supply Unit

In order to power all of the high-torque motors, the MCU, our motor controller, and several LEDs, we needed a high-wattage power supply. In the interest of time and money, we decided to purchase a 600W power supply instead of trying to develop one from scratch. The 24V 25A PSU should provide us with enough voltage to operate our system components and give us enough current to run all of our motors, even in stall conditions.

I. Custom Voltage Regulator Boards

To step down the PSU's voltage to a usable level, we designed a few regulator boards. With the help of TI WEBench, schematics for a 24V-to-3.3V, 24V-to-5V, and 24V-to-8.4V switching regulator circuits were generated. Each of these regulator designs came with its own quirks and issues that needed to be ironed out to function properly under load. That will be discussed later in this report.

J. Relay/Switch Terminal Board

A switchboard has been added to the system to allow the user to manually turn each regulator on or off. This is to prevent the load of the entire system turning on at once, overloading the power supply. Relays attached to rocker switches on the front of the PCB housing will act as our on/off switches.

K. Custom 3D Printed Gimbal

We designed a 2-axis gimbal that is specific to our use case. The gimbal is split into two main parts: the yoke and the base. The base will be stationary and has the primary function of housing all of our PCBs and electrical components. The yoke holds on to our illumination system and will be moving to track our subject. The yoke runs along several large steel ball bearings that are contained in a groove on the bottom of the yoke and the top of the base. These steel balls act as a thrust bearing, allowing the yoke to pan with minimal resistance. The tilting of the illumination system will be handled using a set of servo motors attached to gears on each lateral side of the yoke.

L. Mobile Application Interface

The mobile application provides the primary user interface for the S.T.A.R. system. Through the app, the user can power the system on or off, switch between automatic and manual operating modes, issue direct movement commands, and request target selection. The mobile application connects to the ESP32, which serves as the user-facing gateway for the system. The interface is designed to support both configuration-level actions and low-latency interactive control during operation.

M. Embedded Communication Interface

The software subsystems are connected through a low-latency communication architecture that supports both embedded coordination and user interaction. Internally, the Raspberry Pi and ESP32 exchange tracking and control data through UART, allowing the vision system to deliver frequent position updates with minimal overhead. Externally, the mobile application communicates with the ESP32 over a wireless link for command input and status feedback. This communication structure enables the embedded controller to remain the central coordination point of the system while preserving a lightweight and responsive control path between perception, decision-making, and motion output.

IV. Software Operation and Control Flow

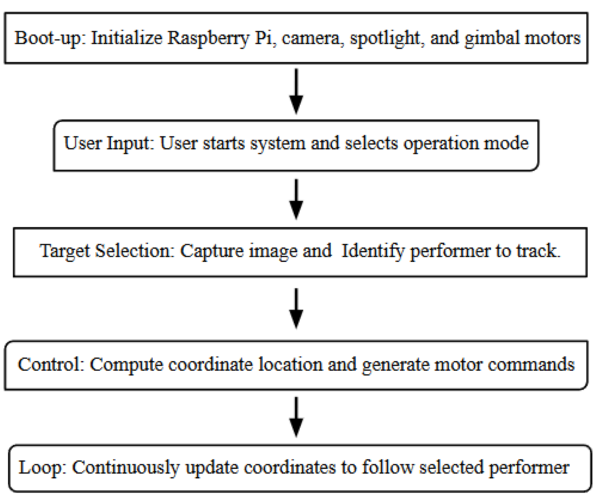

Fig. 5. Software operation and control flow of the S.T.A.R. system. The system initializes, selects a target, computes control commands, and continuously updates tracking in real time

The S.T.A.R. software operates as a continuous perception-and-control loop. During operation, the Raspberry Pi captures image frames from the onboard camera and applies the selected person-detection algorithm to identify the target performer. Once a target is detected, the software computes the target position relative to the center of the image and packages this information as tracking telemetry. The ESP32 receives this telemetry and converts it into motor updates for the pan and tilt axes, allowing the spotlight to follow the performer in real time. This loop repeats continuously while the system remains in autonomous mode.

System operation is organized around explicit software modes to ensure predictable behavior. In idle mode, the system remains powered but does not actively track or respond to motion commands. In autonomous mode, the Raspberry Pi provides live tracking updates and the ESP32 uses those updates to drive the gimbal automatically. In manual mode, tracking-based actuation is suspended and only user-issued control commands are applied to the motors. Separating the system into these operating states prevents conflicts between autonomous and manual behaviors while providing a clear control model for the user.

The control software also incorporates safeguards for state transitions and loss-of-target conditions. When the target is no longer visible or tracking telemetry is delayed beyond an acceptable interval, the ESP32 can stop applying outdated motion updates and transition to a safe behavior. Similarly, when the user switches between manual and autonomous modes, the software prevents abrupt motion by sequencing transitions through the embedded controller rather than allowing concurrent control sources. This state-driven structure improves reliability and makes system behavior more consistent during live demonstrations and performance use.

V. Experiment Setup and Methodology

A. Illumination System

The primary specification that we tested with the illumination system is spotlight uniformity. The goal is to achieve a uniform spot on the wall, without any visible dark spots. The center of the spot should be brighter than the edges, to put emphasis on the subject's face rather than their extremities. Beam and field angles are defined relative to the maximum luminous intensity, typically at 50% and 10% of the peak value, respectively, indicating a center-weighted intensity distribution rather than uniform illumination [1]. In order to test this, the spotlight was aimed at a flat surface and 10 points horizontally were labelled each being 7 inches apart. The same applies to the vertical axis. A basic circuit was implemented that included a photoresistor and a multimeter to measure the current flowing through it. The photoresistor was placed at each point in the spot, allowing relative measurements for each point to be obtained.

B. Voltage Regulators

To test the functionality of our regulator boards, a Keithley 2230-30-1 triple-channel DC power supply, R&S HMC 8012 digital multimeter, and BK Precision 8500B

DC Electronic Load were used. The power supply was set to output 24VDC for all of our regulator testing. The settings used for the DC electronic load were dependent on which regulator we were testing at the time. The multimeter was used as a double check for regulator output voltage.

C. Detection Confidence Score

The human tracking subsystem was tested using the Raspberry Pi CM5, Raspberry Pi HQ Camera, and the Ultralytics YOLO model to verify reliable person detection through the lens system. Live video frames were processed on the Raspberry Pi, and when a person was detected, the model generated a bounding box and confidence score. The center of the bounding box was used to determine the subject’s position in the frame for tracking. To evaluate performance, trials were run where a subject walked in and out of view while detection confidence was recorded. The confidence values were then compared and averaged to measure detection consistency before automated tracking.

D. Tracking Latency

A parameter that was tested was the overall system tracking latency from detection to the movement of S.T.A.R.’s gimbal system. To test this two slow motion (240fps) simultaneous recordings of the gimbal system and the moving subject were recorded and later placed side by side for comparison. By measuring the differences in frames between the subject movement and the response of the gimbal the overall delay of the gimbal can be measured.

VI. Results and Data

A. Illumination System

After measuring the current of the circuit based on the photoresistor at each point for ten trials, the average relative intensity of the spotlight was then plotted as shown in Fig. 6 and Fig. 7, respectively.

Horizontal Photocurrent Distribution

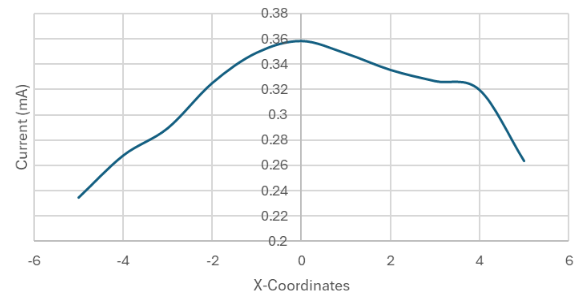

Fig. 6. Plot of the average measured current across 10 trials along the horizontal axis.

Vertical Photocurrent Distribution

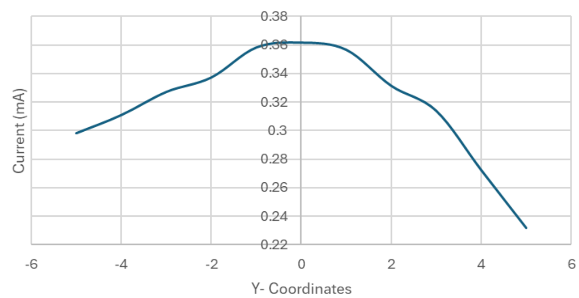

Fig. 7. Plot of the average measured current across 10 trials along the vertical axis.

Taking the average current values of the furthest edge points and comparing them to the center current can provide the percentage of uniformity of the spotlight using the equation below.

$$\frac{Edge\ Photocurrent}{Center\ Photocurrent} * 100 = Uniformity \qquad (1)$$

The results of the calculations are shown in Table I below.

TABLE I
Uniformity Calculations

Edge	Edge Photocurrent (mA)	Center Photocurrent (mA)	Uniformity (%)
Left	0.2348	0.3583	67
Right	0.2637	0.3583	74
Top	0.2317	0.3614	64
Bottom	0.2980	0.3614	82

As expected the illumination system gradually appears dimmer on the outer edges with slower decline on the bottom and right-hand side of the spotlight. A visualization is shown in Fig. 8.

2D Relative Intensity Profile

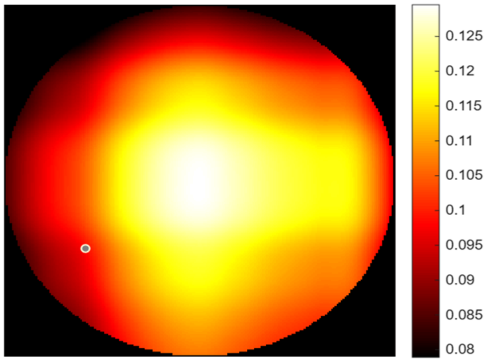

Fig. 8. 2D relative intensity profile from normalization of the photocurrent measurements.

The 2D intensity profile was estimated by combining normalized photocurrent measurements taken along the x and y axes. Since photocurrent is proportional to light intensity, this approach provides a relative visualization of the illumination distribution.

B. Voltage Regulators

Parameter	LMR51450 3.3Vout	LM2576SX 5Vout	LMR51450 8.4Vout
Vout	~3.3V	~5V	~8V
Iout	~0.140A (ESP32 loaded)	~0.50A (0.500A DC Load)	~1.00A (1.00A DC Load)
Vout Ripple	~200 mVpp (unloaded)	~19 mVpp (unloaded)	~100 mVpp (unloaded)
Operating Frequency	~400kHz	~51kHz	~600kHz

C. Detection Confidence Score

Among first detection of a subject onto the tracking software, the initial confidence score was recorded. For each trial, that measurement was plotted with an average confidence between the ten measurements provided as well for comparison. The results are shown in Fig. 9.

Detection Confidence on First Instance

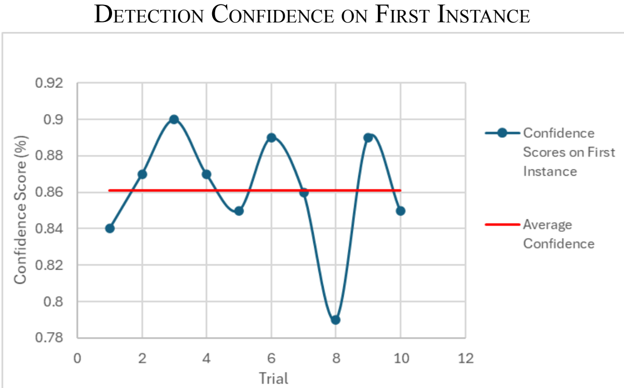

Fig. 9. 2D relative intensity profile from normalization of the photocurrent measurements.

Upon initial recognition of the performer at 7.5m away, the confidence score maintained an average of 86% with a lower threshold of 79%. This result is satisfactory in the ability for S.T.A.R. to be able to identify performers upon them entering the frame.

D. Tracking Latency

When calculating the latency of S.T.A.R., the frame-rate of 240 frames per second in which the slow motion recording was captured proved rather important. Using the following equation, the amount of time taken per each frame can be calculated.

$$\frac{1\ frame}{240\ frames\ per\ second} = 4.17ms \tag{2}$$

From (2), at 4.17ms per frame, the difference in frames (Δf) between the movement of the subject and the movement of the gimbal system when compared side-to-side can be recorded in order to calculate the total latency of S.T.A.R. among 10 different trials.

$$Latency = \Delta f * 4.17ms \tag{3}$$

The recorded data from the side-by-side comparison alongside the calculated latency using (3) was tabulated in Table II below.

TABLE II
Measurement of Latency by Change in Frame

Trial	Change in Frame (Δf)	Latency (ms)
1	68	283.56
2	52	216.84
3	53	221.01
4	62	258.54
5	57	237.69
6	49	204.33
7	53	221.01
8	61	254.37
9	56	233.52
10	58	241.86

Per Table II the calculated latencies performed as expected with a slight delay between tracking to gimbal movement. The results indicate a >300ms response time, which is still quick enough for S.T.A.R. to operate in a real-time stage. As the radius of illumination of the system is quite large, the slight delay of system latency would not interfere with the performance of the system.

VII. Conclusion

The S.T.A.R system successfully demonstrated a compact, autonomous spotlight designed for real-time performer tracking in theatrical environments. By integrating computer vision, motors controls, and an ellipsoidal based illumination system, testing validates the devices ability to detect, track, and illuminate a subject with consistent accuracy. Experimental results confirmed stable tracking performance, responsive gimbal control, and a uniform spot quality at the specified distance. Overall, S.T.A.R meets its objectives and specifications of delivering a stand alone, cost effective solution that simplifies spotlight operation while maintaining reliable and consistent performance.

Acknowledgement

The authors would like to acknowledge Dr. Andrew Klein and Dr. Stephen Eikenberry for their assistance with Zemax simulations and valuable contributions to the imaging lens design.

References

[1] Illuminating Engineering Society, Lighting Handbook, 10th ed. New York, NY, USA: Illuminating Engineering Society, 2011.

[2] J. A. Brons, K. A. Houser, M. Royer, and Y. Ohno, “ANSI/IES TM-30-20: Overview and guidance,” LEUKOS, vol. 17, no. 3, pp. 187–206, 2021.