JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering

Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

Autonomous Camera Tracking System Using Image Processing for Dynamic Educational Content Creation

Valina Sinka[1], Sri Gunawan[2], Muhammad Mustajib[3], Muhammad Rafi Solakhudin[4]
[1,2,3]Management Department, Airlangga University, Surabaya 60115, Indonesia
[4]Automation Engineering, Shipbuilding Institute of Polytechnic Surabaya, Surabaya 60111, Indonesia

ARTICLE INFO

Article historys:

Received : 10/02/2025
Revised : 17/02/2025
Accepted : 30/03/2025

Keywords:

Camera Tracking; Image Processing; MediaPipe; Online Learning Activity

ABSTRACT

In the context of the Online Learning activity, where video content plays a crucial role in educational materials, the demand for effective video production systems has become essential. Traditionally, at least two people are needed to operate cameras, which poses a challenge due to limited human resources. This study addresses this issue by developing a Camera Position Tracking System using image processing, specifically utilizing the MediaPipe framework for real-time tracking of presenters. The system's mechanics enable a DSLR camera to automatically adjust its position based on the presenter's movement, detected within a range of 1.5 to 8 meters. The light intensity required for optimal operation is between 125 and 190 lux. The system's success lies in converting detected position data into motor stepper pulses that move the camera, ensuring efficient, cost-effective video production with minimal human intervention.

Corresponding Author:

Valina Sinka
Management Department, Airlangga University, Surabaya 60115, Indonesia
Email: valinasinka28@gmail.com

1. INTRODUCTION

The increasing reliance on video-based education has revolutionized the way knowledge is delivered, particularly in online learning environments[1]. High-quality educational videos enhance engagement, improve comprehension, and provide flexibility for both instructors and students. However, traditional video production methods require multiple camera operators, leading to increased costs and inefficiencies[2]. Manual camera handling also introduces inconsistencies in framing, limiting the overall quality of educational content.

To address these challenges, this study proposes a autonomous camera tracking system using image processing for dynamic educational content creation[3]. By leveraging image processing techniques and artificial intelligence, the system can autonomously track and follow the movement of presenters without the need for human intervention. Utilizing the MediaPipe framework, the system processes real-time video data to detect and track human positions, ensuring smooth and professional video capture[4]. This innovation is particularly beneficial for educational institutions, online academies, and training centers, where high-quality instructional videos are essential yet often constrained by budget and manpower[5].

The integration of automation in video production not only reduces costs but also enhances efficiency by allowing educators to focus solely on delivering content rather than managing camera operations[6]. With real-time tracking capabilities and adaptive camera positioning, the system ensures that presenters remain in the optimal frame throughout the recording process[7]. This research

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

contributes to the advancement of smart educational technology by providing a scalable, cost-effective, and intelligent solution for modern video-based learning environments[8].

2. RESEARCH METHOD

In This study employs an experimental and system development approach to design and evaluate an autonomous camera tracking system for educational video production[9]. The methodology consists of several key stages, including system design, data collection, system testing, and data analysis

2.1. Hardware System Design

The proposed automated camera tracking system consists of both hardware and software components that work together to enable real-time presenter tracking and dynamic camera positioning. The system is designed to optimize video production for educational purposes by automating camera movement based on the detected position of the presenter[10].

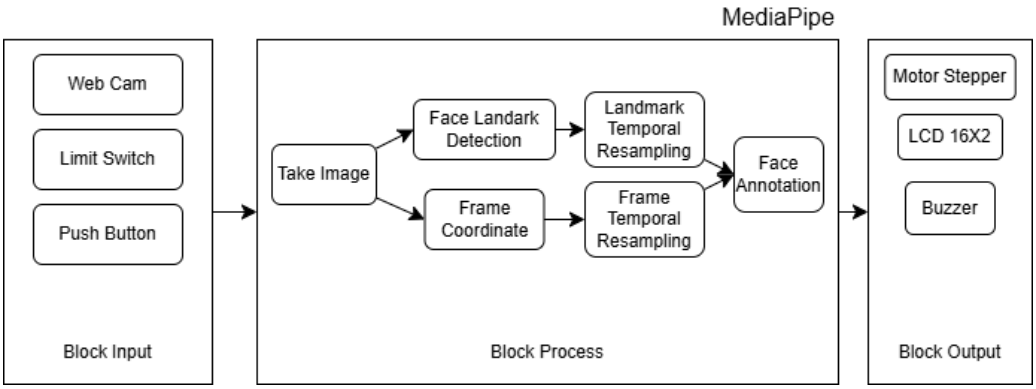

Figure 1 Block Diagram System

The hardware system consists of three main components: imaging devices, a motion control system, and a processing unit. The system employs two cameras a DSLR camera for high-quality video recording and a webcam for real-time presenter tracking. The webcam continuously captures video input, which is processed to determine the presenter's position. To ensure accurate camera movements, a NEMA 17 stepper motor, controlled by a Raspberry Pi 4, is used to adjust the camera's position along two axes (X and Y)[11]. The motor system is integrated with a lead screw mechanism and timing belts, allowing smooth and precise adjustments to keep the presenter within the optimal frame. The Raspberry Pi 4 serves as the central control unit, processing video data and converting detected movements into motor control commands.

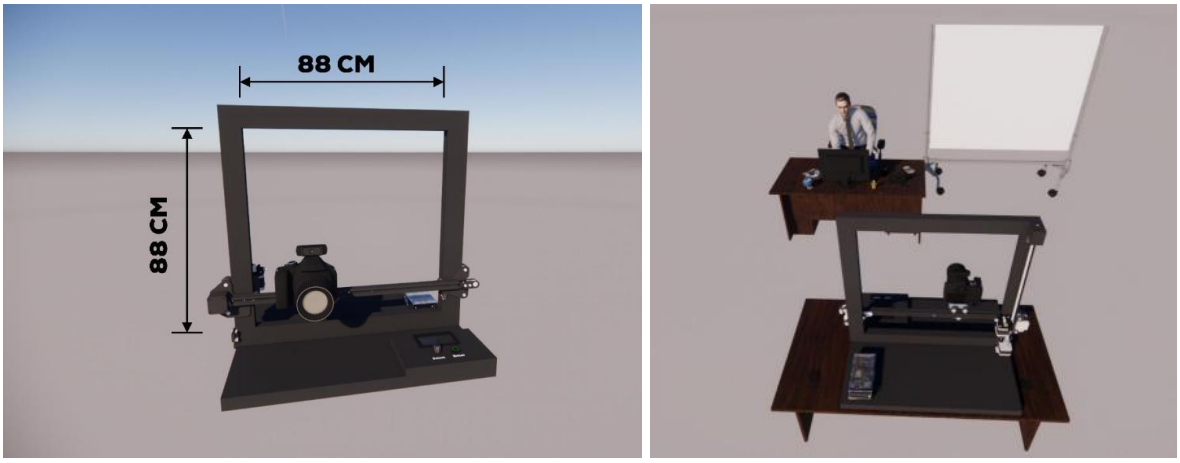

Figure 2 Mechanical Design and Application

JURNAL ECTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

The mechanical design of the autonomous camera tracking system features a rigid aluminum frame with dimensions 88 cm × 88 cm, ensuring stable and precise movement of the mounted DSLR camera. The camera is positioned on a linear rail system driven by a NEMA 17 stepper motor, which enables smooth horizontal tracking based on real-time image processing data. A timing belt and pulley mechanism facilitate motion transmission, ensuring minimal lag in response to presenter movement. This lightweight yet durable structure supports efficient tracking while maintaining stability, making it suitable for dynamic educational content creation[12].

The mechanical system of the automated camera tracking setup relies on precise stepper motor control to ensure smooth movement along both horizontal and vertical axes. The number of pulses required to move the camera is determined by the mechanical components used. For horizontal movement, which utilizes a timing belt and pulley system, the total required pulses are calculated using the formula:

$$Pulse\ max = (belt\ length)/(Pulley\ circumference) \times Pulse\ Stepper \quad (1)$$

for vertical movement, controlled by a lead screw mechanism, the required pulses follow the equation

$$Pulse\ max = (belt\ length\ Screw)/(Pulley\ circumference\ Screw) \times Pulse\ Stepper \quad (2)$$

Given the system specifications, where the timing belt spans 88 cm with a pulley circumference of 4 cm, and the lead screw moves 0.8 cm per revolution, the calculated pulseMax values are 4,400 pulses for horizontal movement and 22,000 pulses for vertical movement. These calculations ensure that the system maintains accurate and stable tracking, allowing the camera to follow the presenter seamlessly[7].

2.2. Software System Design

The software system is responsible for presenter detection, tracking, and camera movement control. The MediaPipe framework is employed for real-time human pose estimation, enabling the system to detect key landmarks on the presenter's body and track movements accurately. The system is programmed using Python and OpenCV, which process video input from the webcam, extract positional data, and translate it into stepper motor commands. The tracking algorithm ensures smooth transitions and minimal delay, optimizing accuracy while preventing sudden or unstable movements[13].

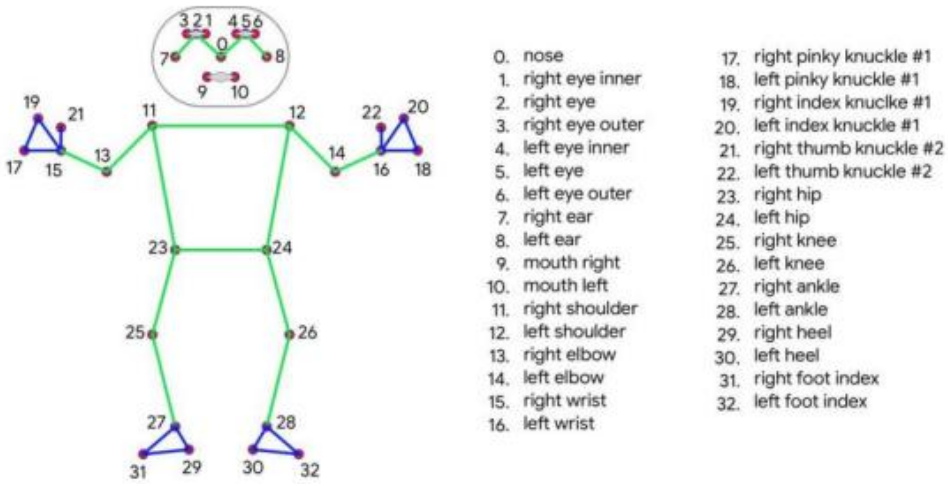

Figure 3 Key Points/Markers of each body part of the MediaPipe Holistic framework

By integrating AI-based image processing with automated hardware control, this system provides an intelligent, cost-effective solution for educational video production. It reduces the need for manual camera operation while improving video quality and consistency, making it highly beneficial for institutions, online learning platforms, and training centers[14].

Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

2.3. Mediapipie Framework

MediaPipe is an open-source framework developed by Google for machine learning-based image processing. This framework is designed to operate in real time with low latency and is compatible with various platforms, including desktop and mobile devices. In this study, MediaPipe Holistic is utilized as it combines face, hand, and body detection models simultaneously to support presenter tracking in educational video production.

MediaPipe employs several key algorithms for image processing. One of the primary models is BlazePose, a Convolutional Neural Network (CNN)-based algorithm capable of recognizing 33 keypoints of the human body in 3D with high accuracy. Additionally, the system incorporates Face Mesh Detection, which uses landmark regression techniques to detect up to 468 facial keypoints, and the Hand Tracking Model, which operates with a multi-stage pipeline to accurately recognize hand and finger positions. In this study, MediaPipe Holistic leverages these techniques to ensure stable and accurate presenter tracking under various lighting conditions and camera angles.

Once the image is captured by the camera, MediaPipe applies pose estimation techniques to analyze the presenter's position and movement. The process begins with pose detection, where the system identifies the presence of a human figure in the frame using deep learning-based models. Subsequently, the system performs keypoint localization, mapping the body's landmarks using Heatmap Regression techniques. To maintain tracking stability, the system integrates the Kalman Filter, which predicts the presenter’s position based on previous data, reducing noise and enhancing movement accuracy. The resulting coordinate data is then converted into X, Y, and Z references, which serve as input for automatic camera movement control.

In this system’s implementation, MediaPipe processes data from a webcam through several key stages. First, the camera captures images at a resolution of 1280×720 pixels. The system then extracts the presenter's keypoints using the BlazePose model. Next, the body’s coordinates are analyzed to determine movement direction. Finally, the extracted coordinate data is converted into stepper motor control commands via a Raspberry Pi 4, allowing the camera to automatically follow the presenter and ensure optimal framing throughout the recording session.

2.4. Data Collection

To evaluate the performance of the automated camera tracking system, data was collected through a series of controlled experiments focusing on detection range, lighting conditions, and tracking accuracy. The detection range was tested by placing the presenter at varying distances from the camera, ranging from 1.5 meters to 8 meters, to determine the system’s optimal operational range. Additionally, lighting conditions were analyzed by adjusting the intensity between 125 to 190 lux, as this range was identified as ideal for maintaining accurate detection and reducing errors in tracking. The system’s tracking accuracy was evaluated by monitoring its responsiveness to different movement speeds and directions of the presenter. The ability of the system to consistently follow the presenter without lag or misalignment was assessed to ensure reliable real-time tracking. These experiments provided crucial insights into the system’s effectiveness and helped refine its ability to maintain smooth, automatic adjustments in educational video production[15].

3. RESULTS AND DISCUSSION

To evaluate the performance of the Autonomous Camera Tracking System, a series of tests were conducted focusing on sensor accuracy, motion detection, tracking responsiveness, framework efficiency, and training improvements. The results are categorized into multiple sections for a detailed analysis.

3.1. System Performance Evaluation

The automated camera tracking system was tested under various conditions to evaluate its overall performance in tracking a presenter for educational video production. The evaluation focused on real-time detection accuracy, response time, motor control precision, and system stability. The system was deployed in an indoor environment with controlled lighting conditions ranging between 125 and 190 lux, as this range was identified as optimal for accurate tracking. The presenter moved within a 1.5 to 8-meter range, and the system successfully adjusted the camera's position to maintain proper framing.

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

Figure 4 Mechanical of the Automatic Camera System

Figure 4 illustrates the mechanical structure of an automatic camera tracking system designed for presenter tracking. The system features a horizontal rail that allows the camera to move autonomously, enabling smooth tracking of the presenter within a designated area. Stepper motors are likely employed to control the camera's movement with high precision. This structure ensures stability and smooth motion, which is crucial for maintaining clear and consistent video recording

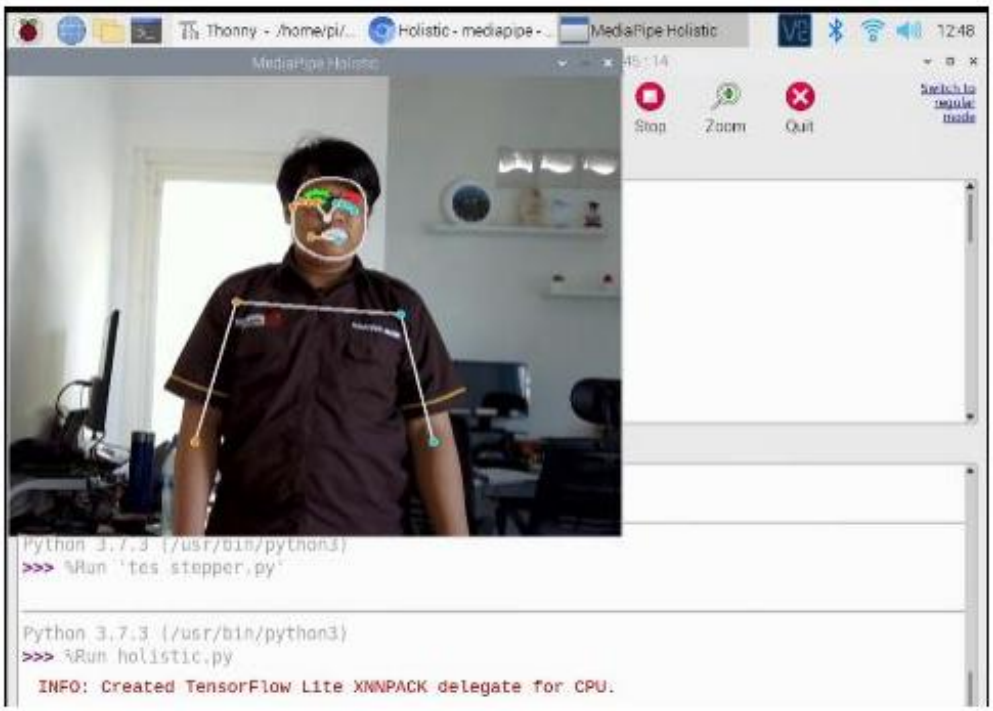

Figure 5 MediaPipe Implementation for Presenter Tracking

Figure 5 demonstrates the MediaPipe Holistic framework in action, effectively recognizing and tracking the presenter's movement in real-time. During testing, the system exhibited high accuracy and minimal lag, ensuring precise detection of facial and body landmarks. The stepper motor's precision was evaluated to assess the system's ability to adjust the camera position dynamically. The camera smoothly followed the presenter along both horizontal and vertical axes, keeping the subject centered within the frame. Additionally, response time analysis revealed that the system quickly detected movement and adjusted accordingly, maintaining stable and seamless tracking. These results highlight the system's reliability in automated educational video production, offering a real-time, AI-driven solution for intelligent camera control.

JURNAL ECTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

3.2. Detection Accuracy Analysis

The accuracy of the automated camera tracking system was evaluated based on its ability to detect and track the presenter at varying distances, angles, and lighting conditions. The system's performance was tested within a range of 1.5 to 8 meters, where accuracy gradually declined as distance increased. At 1.5 to 4.5 meters, the system achieved an accuracy of 98–93%, ensuring stable tracking and precise framing. However, at 6 meters and beyond, accuracy decreased, reaching 85% at 7.5 meters and at 8 meters. This drop is attributed to lower resolution of the detected features and increased background noise, which affects the system's ability to maintain reliable tracking.

Table 1 Detection Accuracy Based on Distance

Distance (m)	Detection Accuracy (%)	Remarks
1,5	98	Optimal detection, clear landmarks
3	96	High accuracy, stable tracking
4,5	93	Slight decrease in accuracy
6	90	Moderate tracking accuracy
7,5	85	Still detectable, stable tracking
8	85	Presenter fully detected

Figure 6 Framework MediaPipie Distance Testing

In addition to distance, the angle of detection also played a significant role in accuracy. The system performed best when the presenter was facing the camera directly, but accuracy slightly decreased when the presenter turned 30° to 60° from the camera, as some facial and body landmarks became partially obscured. At extreme angles (above 75°), tracking errors increased due to occlusions and incomplete landmark detection.

Table 2 Detection Accuracy Based on Angle Presenter

Presenter Angle (°)	Detection Accuracy (%)	Remarks
0° (Front-facing)	98	Ideal detection, all landmarks visible
30°	95	Slight accuracy drop, minor occlusion
60°	88	Partial occlusion of facial landmarks
75°	80	Significant tracking loss in side view
90° (Side view)	70	Poor detection, many landmarks missing

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

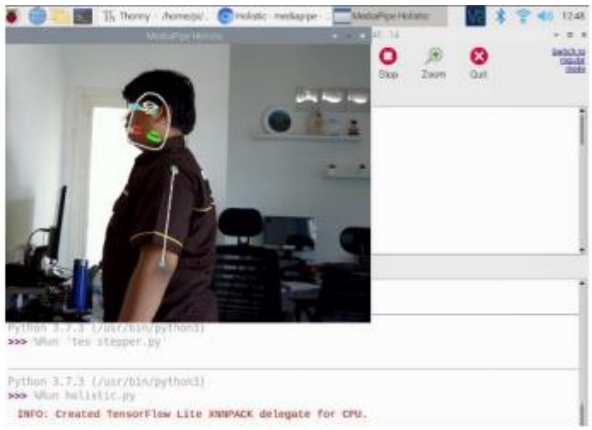

Figure 7 Angle Testing

Lighting conditions also impacted detection accuracy. The system was tested under different illumination levels, where the optimal detection range was observed between 125 to 190 lux. At lower light levels (below 100 lux), tracking performance degraded, leading to occasional misalignment in framing. Conversely, excessive brightness (above 250 lux) caused overexposure, reducing the contrast needed for effective landmark detection. These results indicate that maintaining a well-lit environment within the recommended lux range is essential for ensuring optimal tracking performance.

Table 3 Detection Accuracy Light Intensity

Light Intensity (lux)	Detection Accuracy (%)	Remarks
50	72	Too dark, poor landmark visibility
100	85	Low brightness, minor tracking errors
125	93	Optimal range, stable tracking
150	96	Optimal range, high accuracy
190	98	Ideal lighting conditions

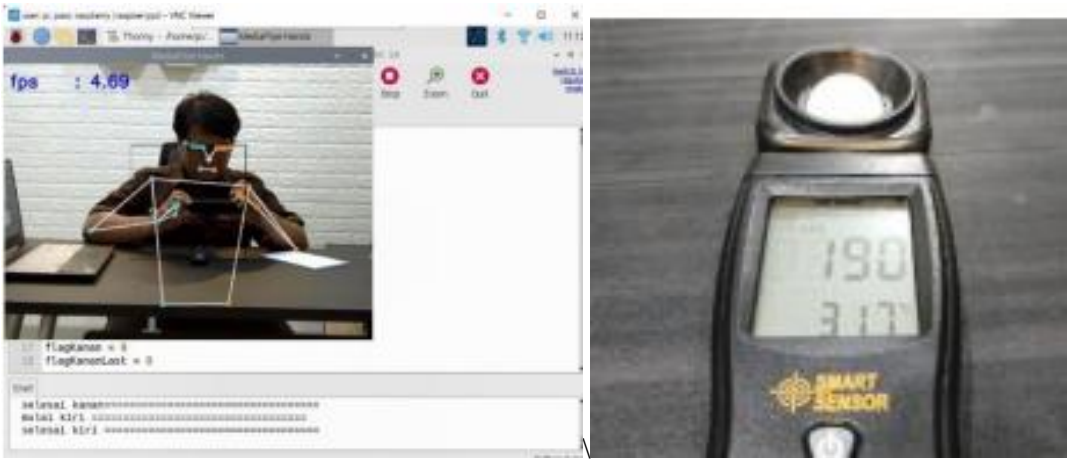

Figure 8 Light Intensity Testing

Overall, the system demonstrated high accuracy within ideal distance, angle, and lighting conditions, confirming its effectiveness for automated educational video production. Future improvements, such as adaptive lighting adjustments and multi-angle tracking, could further enhance its robustness in diverse recording environments.

3.3. Response Time and Processing Speed

The performance of the automated camera tracking system relies on two critical factors: response time and image processing speed, both of which determine the system's ability to track a presenter smoothly in real-time. The response time was measured as the delay between the presenter's movement and the corresponding camera adjustment, ensuring that the tracking remained fluid and accurate. Based

Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

on experimental results, the system achieved an average response time of 253.1 milliseconds, which is fast enough to maintain stable tracking. However, response time tends to increase when the presenter moves beyond 6 meters, where detection becomes less reliable due to reduced image resolution and increased background noise.

In terms of processing speed, the system's frame rate (FPS) was tested under different lighting conditions to evaluate its ability to process and update tracking data in real time. Under optimal lighting conditions (125–190 lux), the system operated efficiently at an average of 30 FPS, ensuring smooth and continuous detection of the presenter's movement. However, when exposed to low-light conditions (below 100 lux), the frame rate dropped to 25 FPS, resulting in slight tracking delays. This reduction was caused by increased noise in the video feed, requiring more computational time for the MediaPipe framework to correctly identify key landmarks.

To further refine the system's real-time tracking, the motor response delay was tested to determine the ideal time interval for stepper motor adjustments. The results showed that the optimal delay for motor movement ranged between 0.005 and 0.1 seconds, providing a balance between accuracy and movement smoothness. If the delay was too low, the motor movements became jerky, whereas a delay that was too high introduced noticeable lag in tracking.

Figure 9 Sample Testing Respons Time with 2,78 Second

Overall, the system demonstrated fast response times and reliable processing speeds, making it suitable for automated educational video production. The findings suggest that maintaining optimal lighting conditions and presenter positioning within 1.5 to 6 meters results in the best tracking performance. Future improvements could focus on hardware acceleration, optimized tracking algorithms, and adaptive lighting compensation to further reduce latency and improve real-time tracking accuracy.

3.4. Motor Control and Camera Movement Precision

The accuracy and stability of the automated camera tracking system rely heavily on the motor control system that adjusts the camera's position in response to the presenter's movements. The system uses a NEMA 17 stepper motor with a lead screw mechanism for vertical movement and a timing belt with a pulley system for horizontal movement. The stepper motor is controlled by the Raspberry Pi 4, which translates tracking data from the MediaPipe framework into precise movement commands.

To evaluate camera movement precision, tests were conducted by measuring the accuracy of motor steps per unit movement. The number of pulses required for movement was calculated using the following formulas:

$$\text{Pulse max} = (\text{belt length})/(\text{Pulley circumference}) \times \text{Pulse Stepper} \quad (3)$$

$$\text{Pulse max} = (\text{belt length Screw})/(\text{Pulley circumference Screw}) \times \text{Pulse Stepper} \quad (4)$$

Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

Based on system specifications the horizontal movement used a timing belt system with a pulley circumference of 4 cm and a belt length of 88 cm, resulting in 4,400 pulses for full horizontal travel. The vertical movement used a lead screw with an 8 mm pitch, requiring 22,000 pulses for full travel of 88 cm.

During testing, the system demonstrated smooth and stable camera movements, ensuring that the presenter remained centered in the frame. However, minor vibrations were observed at higher motor speeds, which could be mitigated by optimizing the motor acceleration profile. The stepper motor delay was fine-tuned to 0.005–0.1 seconds, balancing tracking responsiveness and movement smoothness.

The results confirm that the motor control system effectively maintains precise camera positioning, allowing the system to dynamically adjust the frame in real-time without noticeable lag or jitter. Future improvements may include damping mechanisms to reduce vibrations and the use of brushless DC motors for smoother transitions in high-speed tracking scenarios.

3.5. Comparative Analysis with Manual Camera Operation

To evaluate the effectiveness of the automated camera tracking system, a comparative analysis was conducted against manual camera operation. The comparison focuses on tracking accuracy, response time, movement stability, and operational efficiency, particularly in the context of educational video production.

Tracking Accuracy: The automated system, powered by the MediaPipe framework, maintained an accuracy of 93–98% within an optimal tracking range of 1.5 to 4.5 meters. In contrast, manual tracking depends on the operator's skill and reaction time, which can introduce inconsistencies in framing, especially during rapid presenter movements. At distances beyond 6 meters, the automated system showed a slight accuracy decline due to reduced image resolution and tracking stability, whereas a human operator could still make real-time adjustments.

Response Time: The automated system demonstrated an average response time of 253.1 milliseconds, ensuring fast and seamless camera adjustments. In comparison, manual tracking introduces human reaction delays, which typically range from 500 milliseconds to over 1 second, depending on the operator's experience. The automation significantly reduces latency and human error, improving real-time tracking efficiency.

Movement Stability: The automated system, utilizing stepper motors and optimized motion delay (0.005–0.1 seconds), provided smooth and stable camera movements without noticeable jitter. Manual tracking, however, is prone to sudden shifts and inconsistencies, especially during frequent camera adjustments. Although the automated system performed well under standard conditions, minor vibrations at higher tracking speeds suggest that further optimization could improve motion stability.

Operational Efficiency: Unlike manual camera operation, which requires a dedicated operator, the automated system eliminates human dependency, allowing educators to focus entirely on delivering content. This makes it an ideal solution for e-learning platforms, lecture recordings, and self-produced instructional videos. Additionally, automation reduces production costs and labor requirements, making video creation more scalable and accessible for educational institutions.

4. CONCLUSION

This study successfully developed an automated camera tracking system for educational video production, integrating image processing and motorized camera positioning to enhance efficiency and reduce the need for manual operation. By utilizing the MediaPipe framework, the system effectively detects and tracks the presenter's movements in real time, ensuring smooth and professional video framing. The implementation of stepper motors with lead screw and timing belt mechanisms enables precise camera adjustments along the horizontal and vertical axes, maintaining stable and accurate positioning. Through extensive testing, the system demonstrated a high detection accuracy within an optimal range of 1.5 to 8 meters, with ideal lighting conditions between 125 and 190 lux. The stepper motor control algorithm effectively converted position data into movement pulses, ensuring that the camera seamlessly followed the presenter's position. Additionally, performance evaluations showed that the system maintained a fast response time, minimizing tracking delays and enhancing video quality. Overall, this research contributes to the advancement of automated educational content creation,

Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

providing a cost-effective, scalable, and intelligent solution for modern video-based learning environments. Future improvements may include multi-camera tracking, integration with AI-based speech recognition, and enhanced tracking algorithms to further optimize performance and expand its applications in various educational settings.

REFERENCES

[1] N. Kumar *et al.*, "Educational technology and libraries supporting online learning," *AI-Assisted Libr. Reconstr.*, hal. 209–237, 2024, doi: 10.4018/979-8-3693-2782-1.ch012.

[2] N. Peimani dan H. Kamalipour, "Online education and the covid-19 outbreak: A case study of online teaching during lockdown," *Educ. Sci.*, vol. 11, no. 2, hal. 1–16, 2021, doi: 10.3390/educsci11020072.

[3] A. Khumaidi, "Sistem Tracking Posisi Kamera Menggunakan Pengolahan Citra Untuk Pemusatan Posisi Pengambilan Video di Automation Academy," *J. Tek. Elektro dan Komput. TRIAC*, vol. 9, no. 2, hal. 103–108, 2022, doi: 10.21107/triac.v9i2.16021.

[4] U. Anitha, R. Narmadha, D. R. Sumanth, dan D. N. Kumar, "Robust Human Action Recognition System via Image Processing," *Procedia Comput. Sci.*, vol. 167, no. 2019, hal. 870–877, 2020, doi: 10.1016/j.procs.2020.03.426.

[5] W. Rahmaniar dan A. Hernawan, "Real-time human detection using deep learning on embedded platforms: A review," *J. Robot. Control*, vol. 2, no. 6, hal. 462-468Y, 2021, doi: 10.18196/jrc.26123.

[6] Julham Comaro, I. Malik, M. Mesin Produksi dan Perawatan, P. Negeri Sriwijaya, dan J. Teknik Mesin, "Perancangan Dan Pengembangan Alat Uji Tarik Mini Berbasis Arduino Untuk Spesimen Non-Ferro," *Agustus*, vol. 1, no. 1, hal. 2723–3359, 2020, [Daring]. Tersedia pada: http://dx.doi.org/10.5281/zenodo.4540926.

[7] W. Chamorro, J. Andrade-Cetto, dan J. Solà, "High-speed event camera tracking," *31st Br. Mach. Vis. Conf. BMVC 2020*, no. 2, hal. 1–12, 2020.

[8] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, dan M. Grundmann, "BlazePose: On-device Real-time Body Pose tracking," 2020, [Daring]. Tersedia pada: http://arxiv.org/abs/2006.10204.

[9] T. J. Sánchez-Vicinaiz, E. Camacho-Pérez, A. A. Castillo-Atoche, M. Cruz-Fernandez, J. R. García-Martínez, dan J. Rodríguez-Reséndiz, "MediaPipe Frame and Convolutional Neural Networks-Based Fingerspelling Detection in Mexican Sign Language," *Technologies*, vol. 12, no. 8, hal. 1–22, 2024, doi: 10.3390/technologies12080124.

[10] N. H. M. DHUZUKI *et al.*, "Design and Implementation of a Deep Learning Based Hand Gesture Recognition System for Rehabilitation Internet-of-Things (Riot) Environments Using Mediapipe," *IIUM Eng. J.*, vol. 26, no. 1, hal. 353–372, 2025, doi: 10.31436/IIUMEJ.V26I1.3455.

[11] A. Amarudin, D. A. Saputra, dan R. Rubiyah, "Rancang Bangun Alat Pemberi Pakan Ikan Menggunakan Mikrokontroler," *J. Ilm. Mhs. Kendali dan List.*, vol. 1, no. 1, hal. 7–13, 2020, doi: 10.33365/jimel.v1i1.231.

[12] A. D. Agustiani, S. M. Putri, P. Hidayatullah, dan M. R. Sholahuddin, "Penggunaan MediaPipe untuk Pengenalan Gesture Tangan Real-Time dalam Pengendalian Presentasi," vol. 16, no. 2, 2024.

[13] S. Shriram, B. Nagaraj, J. Jaya, S. Shankar, dan P. Ajay, "Deep Learning-Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19 Spread," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/8133076.

JURNAL ECTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 1, April 2025, pp. 54-64
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i1.4545

[14] A. Specker, "ReidTrack : Reid-only Multi-target Multi-camera Tracking," hal. 5442–5452.

[15] R. Hartmann, F. Al MacHot, P. Mahr, dan C. Bobda, "Camera-based system for tracking and position estimation of humans," *2010 Conf. Des. Archit. Signal Image Process. DASIP2010*, no. April 2014, hal. 62–67, 2010, doi: 10.1109/DASIP.2010.5706247.