Chapter 26

Revolutionising Filmmaking: Artificial Intelligence-driven Motion Control with the Cinema Camera Robot

Ishwar Singh[1]*, Reiner Schmidt[2], Sohaib Al-Emara[2], and Anoop Gadhrri[2]

[1]McMaster University, Ontario, Canada
[2]AXIBO Inc, Canada

Abstract

The AI-Cinema Camera Robot (AI-CCR) introduced a transformative approach to motion control in filmmaking by integrating artificial intelligence with robotic camera systems. Developed by AXIBO (McMaster University graduates) in collaboration with McMaster University, AI-CCR addresses limitations in traditional motion control—such as high cost, lack of flexibility, and technical complexity—by offering a low-cost, modular, portable, scalable, and intelligent solution. The system features a high-payload pan-tilt platform, an AI-enabled single-board computer (SBC) with Coral Edge TPU for real-time inference, and software for automation and tracking. PoseNet integration enables markerless actor and object tracking, while compatibility with Unreal Engine and Dragon Frame supports virtual production and stop-motion workflows. Additionally, AI-CCR supports integration with infrared (IR) tracking technologies for precise actor localisation. Using components like the OptiTrack Active Puck and Prime X cameras, the system achieves sub-millimetre accuracy in motion capture. This IR-based solution, currently deployed at Netflix studios, converts real-time location data into pan-tilt-slide commands, enabling seamless synchronisation with AI-CCR for virtual production workflows. AI-CCR's tracker-less operation, gesture control, and remote accessibility empower creators with precise, repeatable cinematic movements. Real-world deployments at Netflix and Apple studios validate its impact, and future developments—including robotic arms and expanded sensor integration—aim to further democratise high-tech filmmaking. AI-CCR is not just a tool but a platform for innovation, redefining autonomous cinematography for creators of all scales.

Keywords: *AI-Cinema Camera Robot, Pan-Tilt, PoseNet, IR Tracking, TensorFlow Lite, Single Board Computer*

Introduction

Modern filmmaking demands precise, repeatable camera movements, especially for visual effects. Traditional motion control systems face challenges such as sensor degradation, synchronisation issues, and high costs, which limit accessibility for smaller studios. These systems can also restrict actor spontaneity and complicate post-production workflows.[1,2] In response, our team developed a scalable,

[1]Åström, K. J. "Limitations on Control System Performance," *European Journal of Control* 6, no. 1 (2000): 2–20. https://doi.org/10.1016/S0947-3580(00)70906-X.
[2]AXIBO, Python-axibo GitHub Repository, 2023, https://github.com/axiboai/python-axibo.

AI-driven motion control solution that tracks actors and objects with high accuracy using a single-axis robotic platform. Designed to overcome technical and creative barriers, this innovation empowers filmmakers with efficient, flexible tools. It's not just a prototype—it's a step toward democratising advanced cinematic technology worldwide.

Development of AI-CCR System

We have designed, developed, manufactured, and software-integrated an Artificial Intelligence (AI) driven motion control system referred to as AI-Cinema Camera Robot (AI-CCR). This system addresses filmmaking challenges by integrating AI-driven automation and precise motion control. It offers backlash-free movements with 0.01-degree accuracy, ensuring repeatability for multiple takes. The system supports tracker-less virtual production, face and object tracking, and gesture control, making it ideal for VFX integration. Works with Unreal Engine (UE4 & UE5) and Dragon Frame, making it ideal for VFX and stop-motion animation. Additionally, it enhances efficiency with quick setup, remote viewing of feeds, and ease of mobility.

System Overview and Components

The AI-CCR system with the integrated software platform is perfect for cinematic productions, product cinematography, time-lapses, and virtual production. Its intelligent automation and remote accessibility make it a game-changer for filmmakers looking to streamline their workflow. This system consists of the four main components: Pan and Tilt (PT) Platform, a Single Board Computer (SBC), Slider and Software, as shown in Figure **1**.

Figure 1: AI-CCR Major Components.

Pan and Tilt (PT) Platform

We set up three goals for the PT platform, designed, manufactured, and assembled at AXIBO's facilities: track a moving subject (e.g., an actor, an object) in real time; provide smooth, precise motion for cinematic quality; and support camera weight and balance.

Its main components are a payload carrier for a cinema camera, a built-in camera for tracking moving objects, a pan motor, a tilt motor, mounting brackets, a frame, control electronics, and a power supply. This PT requires no balance & as mentioned before, can handle payloads up to 10 KG while still providing backlash-free, 0.01 deg accurate moves. Not only is it the most advanced pan-tilt system, but it is also the strongest in its category.

AXIBO Single Board Computer

The heart of the AI-CCR system is the Single Board Computer (SBC) embedded with AI capabilities, which is designed, assembled and tested at the AXIBO facility. It is based on a 6 core CPU with a dedicated advanced neural network processing (Coral Edge TPU coprocessor),[3] for low-power devices, Figure **2**.

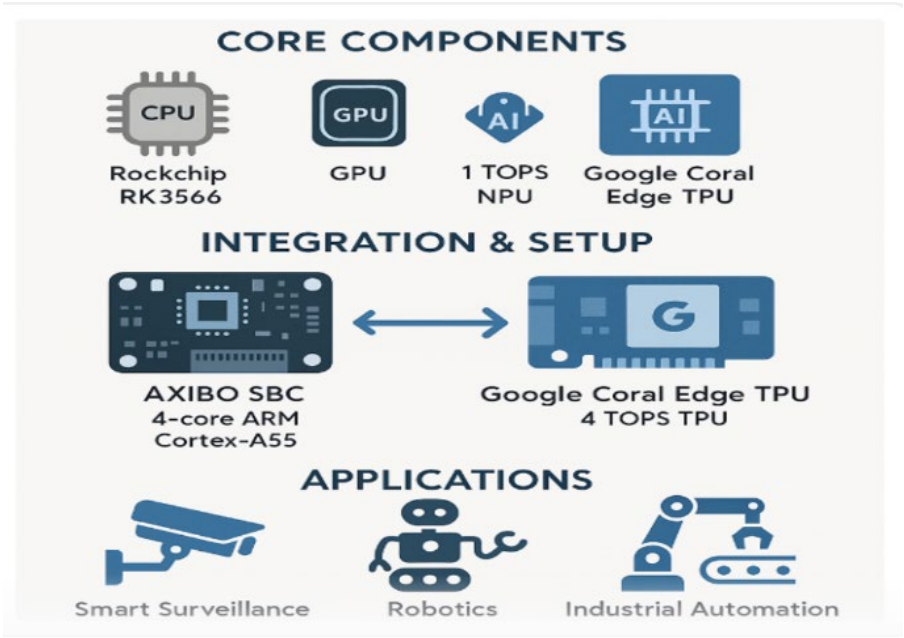

Figure 2: AXIBO Single Board Computer.

The Coral Edge TPU is ideal for real-time, low-power AI inference at the edge. This enables our SBC to be used for the following applications: vision-based object detection; face recognition; pose estimation; license plate recognition; defect detection in manufacturing; smart sensors and meters; predictive maintenance; Intrusion detection; gesture recognition; voice command classification; real-time navigation; obstacle avoidance; and visual SLAM for robotics.

Slider

The AI-CCR slider is a modular, AI-powered motion control system featuring high-speed/ high-torque motors, carbon fibre rails, and joystick control. It supports camera automation and subject tracking and integrates with virtual production tools, offering precise, smooth motion for professional video and robotics applications. The slider, along with PT, makes the AI-CCR a 3 DOF platform.

---

[3]Chen, Yilun, et al., "Adversarial PoseNet: A Structure-Aware Convolutional Network for Human Pose Estimation," In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, (2017): 1212–1221, https://openaccess.thecvf.com/content_iccv_2017/html/Chen_Adversarial_PoseNet_A_ICCV_2017_paper.htm.

AI-CCR Software Embedded on SBC and Film Making

We have developed the following software stack: host CPU runs Linux; edge TPU runtime; TensorFlow Lite; application software. PoseNet enables real-time, markerless human pose estimation by detecting 17 key body points. Integrated with AI-CCR (AXIBO PT4), it allows the robotic camera system to track actors dynamically, adjusting pan, tilt, and slide motions automatically for smooth, cinematic framing.

We have developed a huge volume of applications software for the operation of the AI-CCR and for custom application development, which is available on GitHub. Python-axibo is the library used to interface with AXIBO hardware, including the slider, to control its operations, connect AI-CCR to virtual production tools, as well as write custom applications for AI, motion detection and tracking, stop motion animation and anything else a user may want to use this platform for.[4] AI-CCR can be operated from various sources, such as voice commands, phones, tablets, PCs, and other devices, backed by the integrated Artificial Intelligence (AI) processing unit.

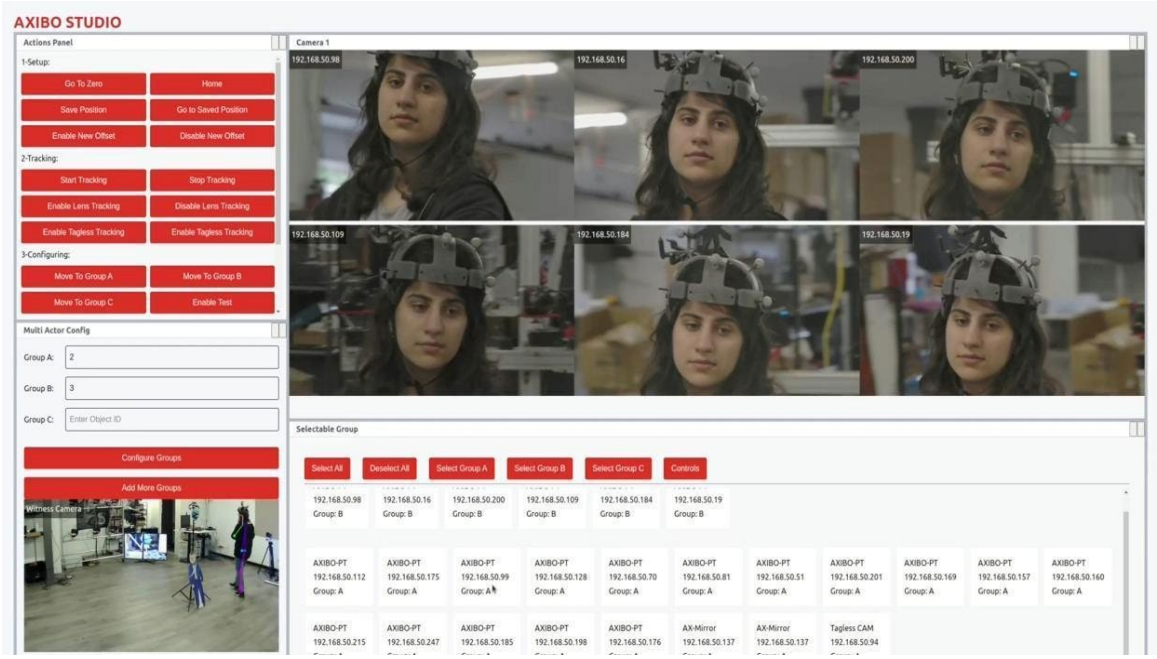

Figure 3: AXIBO Studio Software for Monitoring and Control of Multiple AI-CCRs.

PoseNet Applications In AI-CCR for Actor/Object Tracking

PoseNet[5,6,7] is a lightweight, real-time pose estimation model that detects human body key points, like eyes, ears, nose, shoulders, etc. (Figure **4** and Figure **5**), from images or video. It works by identifying 17 key points on the human body and estimating their positions in 2D space (Figure **6**). An image or

[4]Coral by Google, "Edge TPU Overview," 2020, https://coral.ai/docs/edgetpu/overview/.
[5]Google, "PoseNet: Real-Time Human Pose Estimation," 2018, https://github.com/tensorflow/tfjs-models/tree/master/posenet.
[6]A. A., Khan, et al., "Comparative Analysis of OpenPose, PoseNet, and MoveNet Models for Pose Estimation." *Traitement du Signal,* 39, no. 1 (2022): 111–118. https://doi.org/10.18280/ts.390111.
[7]Mars Motion Control, "The Role of Motion Control Technology in Modern Film and Television Production," n.d. https://marsmoco.com/blog/role-of-motion-control-technology-in-modern-film-and-television-production/.

video frame is input to the SBC (Figure **6**). It provides a set of key points for each detected person, each with (x, y) coordinates and a confidence score. The tracking logic assigns consistent IDs to actors across frames using spatial proximity or tracking algorithms (e.g., SORT, Deep SORT).[8,9]

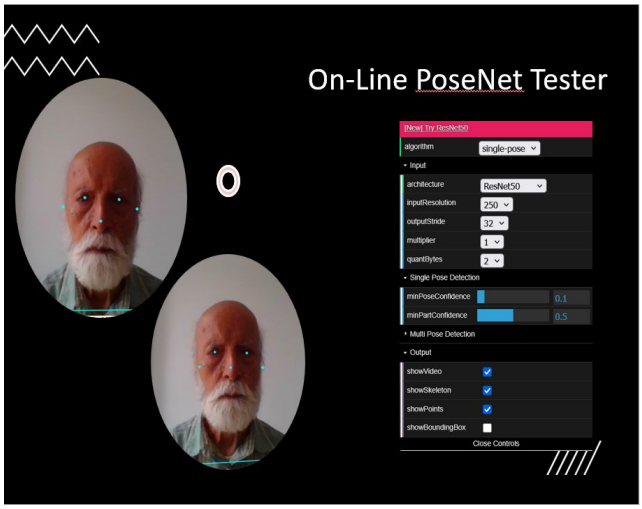

Figure 4: Key Points Selection Example.

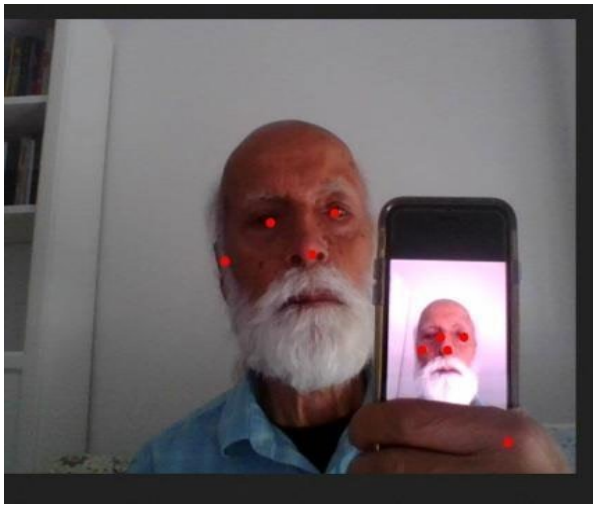

Figure 5: Multiple Actor Key Points.

A key element is training a model for these tasks and deploying it. The model is trained in TensorFlow, converted to TFLite, quantised to INT8 to run on TPU and then compiled with the edge compiler. The camera feed from the built-in camera is processed in SBC with PosNet detecting key points, which are used to move the pan/tilt and slide motors to follow the actor. For example, if a person moves left, the PT can smoothly follows using pose data.

AI-CCR Applications

AI-CCR revolutionises filmmaking by offering intelligent, portable automation tools for solo creators, choreographers, trainers, and virtual production teams (Figure **8**). Designed for real-time performance

[8]Unreal Engine, 2023, "Virtual Production Tools," 2023, https://www.unrealengine.com/en-US/virtual-production.

[9]Wojke, Nicolai, et al., "Deep SORT: Deep Learning for Multiple Object Tracking," In Proceedings of the IEEE International Conference on Image Processing, 2017, https://arxiv.org/abs/1703.07402.

on AXIBO SBC edge devices, it eliminates the need for markers or suits, ensuring privacy with no facial recognition or identity tracking.

Figure 6: Body Points Mapping.

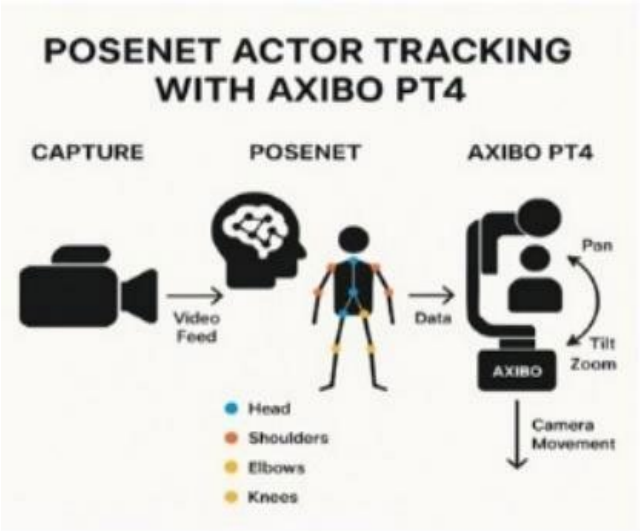

Figure 7: PoseNet Tracking Setup.

Key features include:

- Compatibility with phones, tablets, and PCs
- Ethernet & Wi-Fi for remote control
- Control via app, Xbox controller, or onboard OLED
- Stealth drive for silent, repeatable motion
- Precise, backlash-free movement (0.01°)
- Face, object, and joint tracking without keyframes

AI-CCR auto-detects subjects, syncs with sliders, and operates via any web browser—no downloads required. Multiple devices can be connected and controlled remotely, with live camera feeds and previews enhancing workflow. It supports general object tracking using onboard or mounted cameras and includes one-click parallax for efficient product cinematography. For virtual production, AI-CCR

integrates seamlessly with Unreal Engine (versions 4.27 to 5.1)[10], Figure **9**, supporting FreeD, LiveLink, Aximmetry, and Smode. The Axibo PT motion-controlled slider enables high-speed or slow-motion VFX shots, making advanced robotic filmmaking accessible to creators of all sizes. In addition, one can generate 3D images by converting multiple images from several AI-CCRs to a 3D model using a tool like Meshy AI, using its "Image to 3D" tool, enabling the multi-view option, and uploading your main image plus up to three additional images of the object from different angles. Alternatively, one can use a photogrammetry software like Polycam by uploading all images and selecting the option to create a 3D model from the capture.

Figure 8: AI-CCR Applications.

Beyond automation, the system supports traditional pan-tilt-slide functions like tracking, timelapses, keyframing, astrophotography, and zoom/focus control—ideal for everyday shooting. AXIBO Studio provides unified monitoring and control across platforms and devices.

Figure 9: AI-CCR and Virtual Production Example.

[10]Wojke, Nicolai, et al., "Deep SORT: Deep Learning for Multiple Object Tracking," https://arxiv.org/abs/1703.07402.

Integration with other tracking and location technologies

AI-CCR does not natively advertise integration with other location and tracking technologies, such as GPS, RTK, UWB or IR technologies out of the box; there are ways to incorporate such technologies depending on your setup and goals. However, we have developed and integrated an IR-based solution for tracking actors with the AI-CCR system. This system is currently deployed at the Netflix film-making studio. Basic steps are as follows: IR camera reads real-time location data from an IR module, converts that data into pan/tilt/slide coordinates, and sends those commands to the PT via the SDK. IR Technology for Actor Tracking Infrared (IR) tracking technology enables precise, real-time localisation of actors in motion capture environments. By emitting synchronised IR signals, active tracking devices like the OptiTrack Active Puck can be detected by specialised cameras, such as the OptiTrack Prime X series (Figure **10** and Figure **11**). This system allows for sub-millimetre accuracy in determining an actor's position and orientation, forming the foundation for seamless integration with AI-CCR in virtual production workflows.

Figure 10: IR Tracking Components.

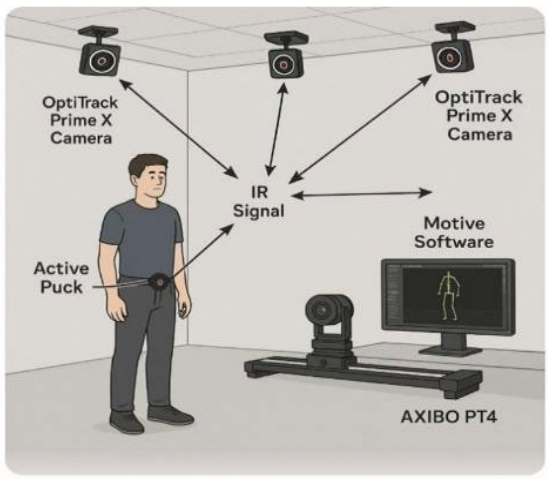

Figure 11: Actor Tracking Setup.

Conclusion

The AI-Cinema Camera Robot (AI-CCR) represents a transformative leap in motion control technology for filmmaking. By integrating advanced AI capabilities such as PoseNet for real-time pose

estimation and Coral Edge TPU for low-power inference, the system delivers unmatched precision, automation, and adaptability. Its modular architecture—including a high-payload pan-tilt platform for filming camera, an intelligent single-board computer with a built-in small camera for PoseNet tracking, and AXIBO source control software—empowers filmmakers to achieve complex shots with minimal setup and maximum creative freedom, at a very low cost.

AI-CCR's compatibility with industry-standard tools like Unreal Engine and Dragon Frame, along with its support for trackerless virtual production and infrared actor tracking, positions it as a versatile solution for both indie creators and major studios. The system's real-time tracking, gesture control, and remote accessibility redefine the boundaries of autonomous cinematography.

Collaborations with institutions like McMaster University and deployments at Netflix and Apple validate its real-world impact. Looking ahead, AXIBO's roadmap—including the E-JIB MINI robotic arm and expanded sensor integration—promises to further democratise high-tech filmmaking.

In essence, AI-CCR is not just a tool—it's a platform for innovation, enabling storytellers to push the limits of visual expression while streamlining production workflows. It marks a pivotal step toward accessible, intelligent, and scalable cinematic technology.

References

Åström, K. J. 2000. "Limitations on Control System Performance." European Journal of Control 6 (1): 2–20. https://doi.org/10.1016/S0947-3580(00)70906-X.

AXIBO. 2023. Python-axibo GitHub Repository. https://github.com/axiboai/python-axibo.

Chen, Yilun, Chunhua Shen, Xiu-Shen Wei, Lingqiao Liu, and Jian Yang. 2017. "Adversarial PoseNet: A Structure-Aware Convolutional Network for Human Pose Estimation." In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1212–1221. https://openaccess.thecvf.com/content_iccv_2017/html/Chen_Adversarial_PoseNet_A_ICCV_2017_paper.html.

Coral by Google. 2020. "Edge TPU Overview." Accessed August 21, 2025. https://coral.ai/docs/edgetpu/overview/.

Google. 2018. "PoseNet: Real-Time Human Pose Estimation." Accessed August 21, 2025. https://github.com/tensorflow/tfjs-models/tree/master/posenet.

Khan, A. A., M. A. Khan, and S. A. Khan. 2022. "Comparative Analysis of OpenPose, PoseNet, and MoveNet Models for Pose Estimation." Traitement du Signal 39 (1): 111–118. https://doi.org/10.18280/ts.390111.

Mars Motion Control. n.d. "The Role of Motion Control Technology in Modern Film and Television Production." Accessed August 21, 2025. https://marsmoco.com/blog/role-of-motion-control-technology-in-modern-film-and-television-production/.

Unreal Engine. 2023. "Virtual Production Tools." Accessed August 21, 2025. https://www.unrealengine.com/en-US/virtual-production.

Wojke, Nicolai, Alex Bewley, and Dietrich Paulus. 2017. "Deep SORT: Deep Learning for Multiple Object Tracking." In Proceedings of the IEEE International Conference on Image Processing. https://arxiv.org/abs/1703.07402