

Automated Camera Tracking Software

Furkan Mucahid YILDIZ¹, Mahmut Kursad ARPACIOGLU¹

*¹Turkish Aerospace Inc. Fethiye Mah. Havacilik Bulvari No:16 Kahramankazan/Ankara/Turkiye
furkanmucahid.yildiz@tai.com.tr, mkarpacioglu@tai.com.tr*

Abstract:

Flight testing, ground-based cameras are used to provide an independent visual reference for validating aircraft motion, system behavior, and test execution in real time. Ground-based camera tracking of airborne platforms is traditionally achieved using image processing techniques, specialized PTZ hardware, or computationally intensive visual tracking algorithms. While effective, such approaches introduce increased system complexity, latency, and cost. In many operational and flight-test scenarios, however, the precise position of the aircraft is already available through real-time GPS telemetry. This study presents a geometry-driven camera tracking system that deliberately avoids image processing and instead relies solely on GPS data to control camera orientation.

The proposed system computes real-time pan and tilt commands by combining the aircraft's GPS position with the known geodetic location of each camera. The architecture supports simultaneous tracking by multiple cameras, where each camera independently calculates its own azimuth and elevation angles based on its fixed position. Camera control is performed using the ONVIF protocol, enabling compatibility with a wide range of standard commercial IP cameras without requiring specialized hardware.

In addition to directional control, the system integrates automatic zoom and focus adjustment to maintain target visibility across varying distances, further enhancing tracking robustness without visual feedback. The resulting solution provides deterministic behavior, low computational load, and high scalability, making it well suited for flight test environments, temporary test ranges, and cost-sensitive surveillance applications.

This study demonstrates that camera tracking problems can be effectively addressed through sensor-based geometric prediction rather than visual perception, offering a practical and efficient alternative for telemetry-driven observation systems.

Key words: Visual aircraft tracking, PTZ camera, ONVIF, flight test instrumentation, telemetry-cued tracking, sensor calibration

Introduction

Maintaining a continuous visual record of an aircraft during flight test activity is a fundamental requirement of any flight test center. Beyond providing photographic documentation of test events, the video feed is used by test engineers in real time to assess attitude, configuration, and the proximity of the aircraft to range boundaries or hazards. Dedicated optical tracking platforms designed for this purpose are highly capable but typically cost in the range of hundreds of thousands of dollars per unit and require significant ground infrastructure, calibration effort, and trained operators.

The system described in this paper takes a different approach. Rather than relying on

dedicated optical trackers, it drives commercial off-the-shelf ONVIF-compatible PTZ surveillance cameras—the same class of cameras already deployed throughout most flight test ranges for general observation—using only the live telemetry stream that is already routed to ground stations during every test. This places the work in the family of telemetry-cued or GPS-aided electro-optical tracking, but at a fraction of the cost and complexity of comparable solutions in the literature.

All target-pointing computations reduce to elementary spherical trigonometry. The system contains no image processing or computer vision component and therefore requires no GPU hardware; it runs comfortably on any modern

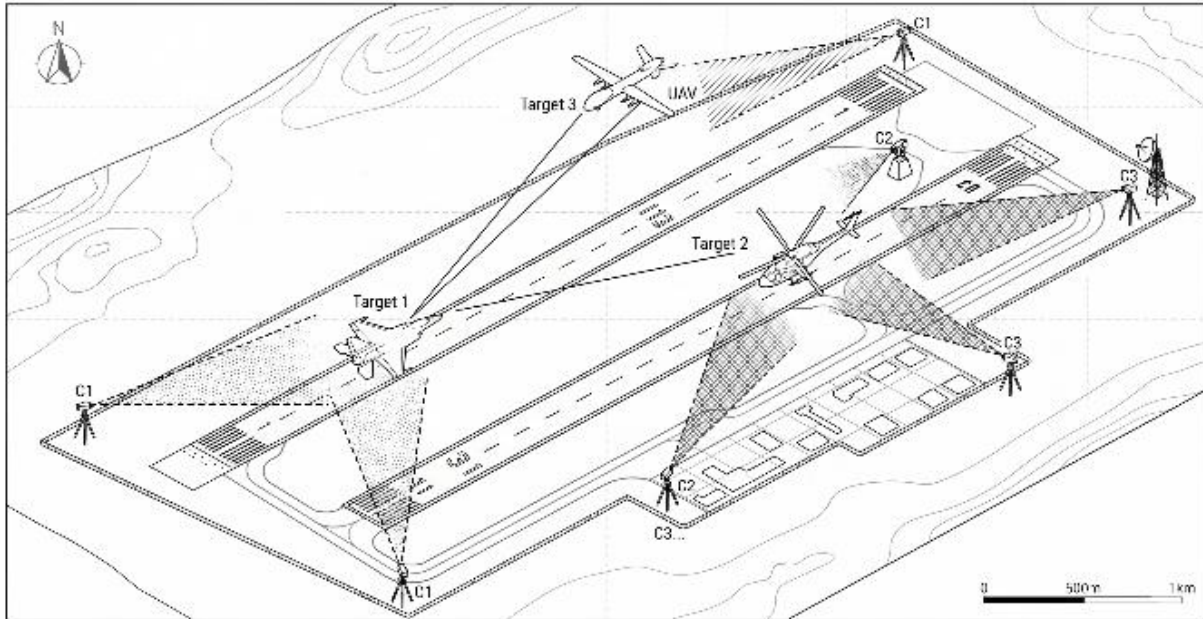


Figure 1: System Deployment Overview

personal computer while controlling eight cameras in parallel. A single coordination process is capable of directing up to several PTZ cameras of differing makes and models toward up to several different aircraft simultaneously, with each camera assignable to any target through the operator interface.

The system has been used operationally to track a broad portfolio of platforms, ranging from high-agility maneuvers of the HÜRJET jet trainer, through long-endurance flights of ANKA and AKSUNGUR unmanned aerial vehicles, autorotation tests of the GÖKBAY and ATAK helicopters, sorties involving KAN. This breadth of platform types—with airspeeds ranging from near-hover to transonic and engagement geometries from a few hundred meters to several kilometers—has provided a demanding validation envelope and exposed several mechanical and software failure modes that have since been addressed by the calibration and prediction subsystems described below.

The principal contributions of this work are: (i) a complete telemetry-cued multi-camera tracking architecture that does not depend on image processing; (ii) a per-camera sinusoidal regression scheme that compensates for mechanical mounting errors of the PTZ gimbal; (iii) a logistic-regression model of the distance-to-focus relationship that enables open-loop autofocus on low-contrast targets such as aircraft against the sky; and (iv) a hybrid command strategy that uses ONVIF for portability and vendor-specific APIs for advanced control.

System Architecture

The system consists of a single central computer connected over Ethernet to a number of distributed PTZ cameras and to the telemetry source. The software is organized as six functional layers, depicted in Figure 1: data acquisition, geometric solver, automatic zoom and focus, calibration and correction, command and protocol, and the operator interface.

The data acquisition layer subscribes to the telemetry stream and validates each incoming sample. The geometric solver layer converts the aircraft position into the pan and tilt angles each camera must command. The automatic zoom and focus layer computes the appropriate optical settings for the current engagement geometry. The calibration and correction layer applies per-camera offsets, zoom-dependent corrections, and the sinusoidal tilt regression that accounts for mechanical misalignment. The command and protocol layer translates the resulting commands into ONVIF and vendor API messages and dispatches them to the cameras. Finally, the operator interface layer exposes target assignment, manual offset tuning, and the on-screen display (OSD) overlays that keep the operator informed of system state.

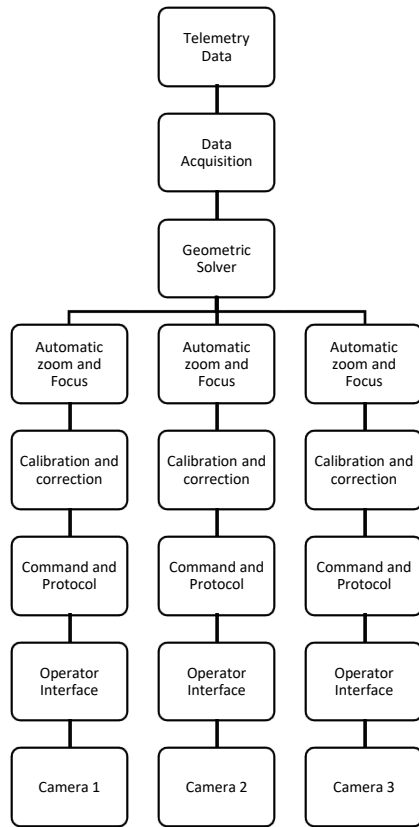


Figure 2: Six-layer system architecture. Telemetry, RTK-derived camera positions, and operator inputs flow through the processing layers; commands are dispatched to heterogeneous PTZ cameras over ONVIF and LAPI

Telemetry Data

The system is fed by the same telemetry stream that ground stations use during flight test. Latitude, longitude, altitude, heading, and ground speed are sampled at 10 Hz, which is sufficient to follow the angular dynamics of any aircraft in the supported portfolio without aliasing. Each sample is parsed and passed to the position-processing pipeline.

Telemetry health is monitored continuously, both from the telemetry status indicator by the ground station and from the sample-to-sample consistency of the data itself. A sample is rejected if any of the following thresholds are exceeded with respect to the previous good sample: a latitude change of more than 0.006 degrees (approximately 670 m), a longitude change of more than 0.008 degrees, an altitude change of more than 666 m, or telemetry status indicators health level. When a sample is rejected, the cameras continue to be commanded with the most recent valid position rather than being driven by the corrupted one. This prevents brief telemetry spikes—typically

caused by datalink dropouts or GPS multipath—from throwing the cameras off the target.

Precise Camera Positioning

The accuracy of the final pointing solution is bounded by how well the camera positions themselves are known. Each camera in the system is surveyed using RTK GNSS equipment, yielding centimetre-level absolute positions in WGS-84. These coordinates, together with the camera mounting altitude, are stored in the system configuration and are used as the origin of every geometric calculation. Any error in the camera position propagates directly to the pan and tilt commands, so this step is treated as part of the installation procedure rather than an operational task.

Target Position Processing and Geometric Solver

Camera commanding is done in absolute pan, tilt, and zoom coordinates rather than continuous velocity commands, which avoids drift accumulation and allows the commanded position to be recomputed independently on every cycle. The required pan and tilt angles are obtained from the relative geometry between the camera and the aircraft, both expressed in WGS-84 latitude, longitude, and altitude:

$$\text{Camera } (\varphi_c, \gamma_c, h_c)$$

$$\text{Aircraft } (\varphi_a, \gamma_a, h_a)$$

The horizontal great-circle distance between the two points is computed by the haversine formula, which is numerically stable for the short to medium ranges typical of flight test engagements:

$$a = \sin^2\left(\frac{\varphi_a - \varphi_c}{2}\right) + \cos(\varphi_c) \cos(\varphi_a) \sin^2\left(\frac{\gamma_a - \gamma_c}{2}\right)$$

$$d = 2R \cdot \arcsin(\sqrt{a}) \quad (R \cong 6371 \text{ km})$$

$$\Delta h = h_a - h_c$$

$$D = \sqrt{d^2 + \Delta h^2}$$

The azimuth (pan) angle is computed from the standard great-circle bearing equation. The use of atan2 rather than a single-argument arctangent is essential here, since it preserves the correct quadrant across the full 0–360° range:

$$\theta_{az} = \text{atan2}(\sin(\Delta\gamma) \cos(\varphi_a) \cos(\varphi_c) \sin(\varphi_a) - \sin(\varphi_c) \cos(\varphi_a) \cos(\Delta\gamma))$$

The elevation (tilt) angle is the simple ratio of altitude difference to horizontal ground distance. The ground distance d is used in place of the slant range D because the gimbal measures tilt with respect to the local horizontal, not to the line of sight:

$$\theta_{el} = \arctan\left(\frac{h_a - h_c}{d}\right)$$

Pan is referenced to true north (0° = north, increasing clockwise). The slant range D is retained for the autozoom and autofocus subsystems described in the next section.

Latency Compensation and Position Prediction

Because the telemetry pipeline introduces a measurable end-to-end latency, commanding the cameras with the last received position would leave the aircraft systematically lagging in the frame, particularly during high-speed segments. To compensate, the most recent valid position is propagated forward by the measured pipeline latency using the standard great-circle destination equations:

$$pred_\varphi = \arcsin\left(\sin(\varphi) \cos\left(\frac{d}{R}\right) + \cos(\varphi) \sin\left(\frac{d}{R}\right) \cos(\psi)\right)$$

$$pred_\gamma = \gamma + \text{atan2}\left(\sin(\varphi) \sin\left(\frac{d}{R}\right) \cos(\psi), \cos\left(\frac{d}{R}\right) - \sin(\varphi) \sin(pred_\varphi)\right)$$

Here $d = v \cdot \Delta t$ is the predicted ground-track distance, where v is the ground speed and Δt the prediction horizon configured to match the measured pipeline latency; $R \approx 6371$ km is the Earth radius; φ and γ are the latest valid latitude and longitude; and ψ is the aircraft heading from the telemetry stream. Altitude is propagated linearly from the most recent climb or descent rate.

Automatic Zoom

The autozoom logic is designed so that the platform with the largest wingspan in the supported portfolio still fits comfortably within the camera frame at any operational distance, and so that smaller platforms are progressively zoomed in as range increases. The required horizontal field of view (HFOV) to image a target of half-width L at slant range D is

$$Required_{HFOV} = 2 \arctan(L/D)$$

The corresponding zoom command, expressed in the $[0, 1]$ normalised range that the cameras accept, is the ratio of the required HFOV to the camera maximum HFOV. For the camera class used in the system, the maximum HFOV at full wide is approximately 38.4° ; L is taken as half the wingspan of the largest supported platform. The relationship is therefore:

$$Zoom = HFOV_{cam} / (2 * \arctan\left(\text{desired length} \cdot \frac{180}{\pi}\right) / D)$$

Automatic focus

Open-loop autofocus is used in preference to the contrast-detection autofocus built into the cameras, because the typical scene—an aircraft against a uniform sky—provides too little contrast for the camera firmware to lock reliably. Instead, the system commands an absolute

focus position derived directly from the known target distance.

The mapping from distance to optimal focus value was characterised empirically. For each camera, between fifteen and twenty manual focus adjustments were performed at distances spanning approximately 100 m to 3500 m, and the resulting (distance, focus) pairs were logged. The data points exhibited a clear saturating shape consistent with the optical behaviour of a zoom lens: focus changes rapidly at short range and asymptotes toward a fixed value as the target approaches optical infinity. A four-parameter logistic (sigmoid) function was therefore fitted to the data using non-linear least squares:

$$\text{Focus (logistic regression)} = \frac{L}{1 + e^{(-k(d+d_0)-x_0)}}$$

where L is the upper limit of the focus motor (9350 in the cameras used), k controls the steepness of the curve, d_0 is a distance offset, and x_0 sets the inflection point of the sigmoid. For a representative camera the fitted parameters were $L = 9350$, $k = 0.00284$, $d_0 = 25$ m, and $x_0 = 2.927$, placing the inflection point at approximately 1000 m. The fit is illustrated in Figure 2.

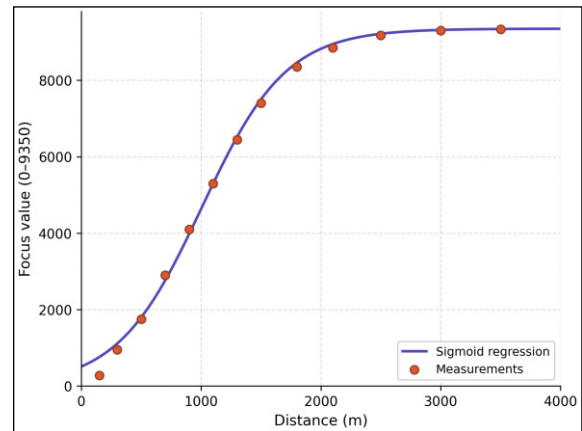


Figure 3: Distance-to-focus relationship for a representative camera. Manual measurement points are shown together with the fitted four-parameter sigmoid function used at run time to command absolute focus positions.

At run time the focus value is recomputed on every camera update and dispatched as an absolute focus command. Because the model is purely a function of distance, the focus tracks the target through approach and recession phases without any image-based search, eliminating the focus-hunting artefacts common to contrast autofocus on aircraft.

Calibration and Correction

Mechanical alignment errors introduced during PTZ camera installation—particularly any tilt of

the gimbal base relative to the local horizontal—cause systematic pointing errors that grow with zoom level and can take the aircraft completely out of frame at high magnification. Even a tilt-axis misalignment of half a degree, which is invisible to the naked eye and well within typical mounting tolerances, is sufficient to displace the target by several frame heights at full zoom. These errors are characterised and removed by a per-camera regression model.

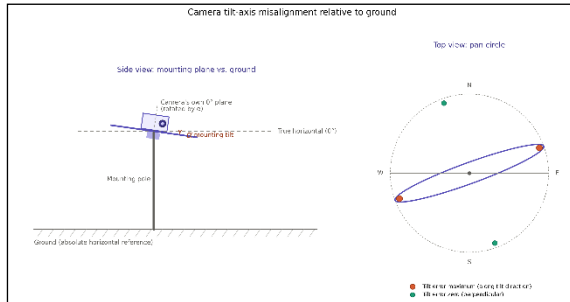


Figure 4: Camera tilt-axis misalignment relative to ground

A set of fixed survey markers is placed in the surroundings of each camera and their positions measured with the same RTK GNSS equipment used for the camera positions themselves. A typical calibration uses between ten and fifteen markers spread evenly around the camera in azimuth, at a variety of distances and elevations. For each marker the geometric solver computes the pan and tilt angles that the camera would need to command if it were perfectly aligned. The camera is then commanded to those angles, and an operator records the residual offset required to bring the marker to the centre of the image. The difference between the actual tilt achieved and the geometrically expected tilt is the tilt error ϵ , which is the quantity to be modelled.

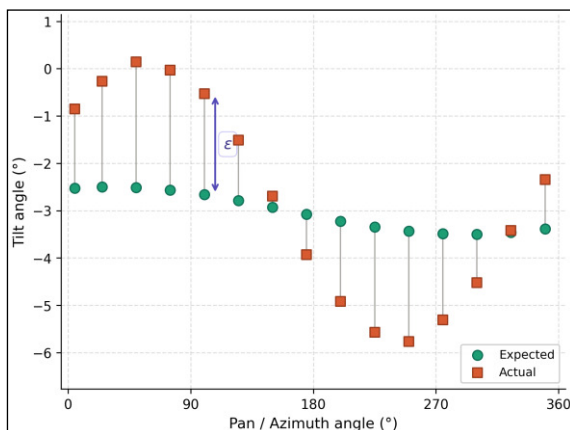


Figure 5: Tilt error measurement. For each survey marker the expected tilt (green) and the actually achieved tilt (orange) are recorded across the full pan range. The difference between the two values constitutes the per-azimuth tilt error ϵ .

When the tilt error is plotted as a function of the pan angle, the resulting curve is sinusoidal. This shape has a direct physical interpretation: if the tilt axis of the gimbal is mechanically inclined by a small angle α relative to the local horizontal, then as the gimbal pans through 360° the angle between the optical axis and the true horizontal varies as $\alpha \cdot \sin(p - p_0)$, where p_0 is the azimuth at which the inclination is maximal. Real installations also exhibit a constant offset that absorbs zero-position errors in the tilt encoder, leading to the four-parameter form:

$$\epsilon(p) = A \sin(Bp + C) + D$$

Here A is the amplitude of the misalignment (typically in the range $1\text{--}3^\circ$ for hand-installed cameras), B fixes the period to one full cycle per pan revolution ($B \approx 2\pi/360 \approx 0.01745$ rad/deg when p is in degrees), C is the phase shift identifying the worst-case azimuth, and D is the constant offset. The four parameters are fitted to the per-marker error measurements separately for each camera. The result for a representative camera is shown in Figure 6.

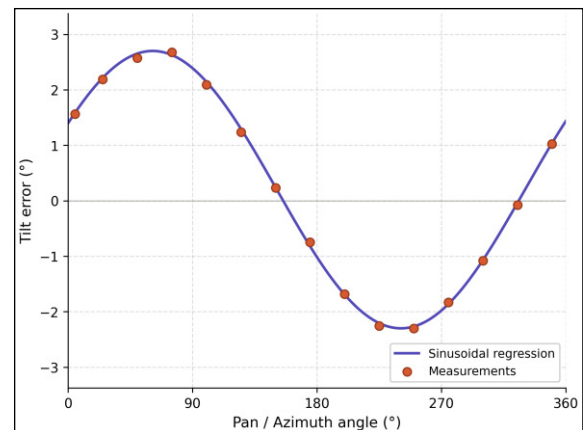


Figure 6: Sinusoidal regression fitted to the tilt-error measurements. The fitted function $\epsilon(p)$ is subtracted from the geometric tilt command at run time to compensate the gimbal misalignment.

At run time the calibration model is evaluated for the current pan angle and subtracted from the geometric tilt before the command is dispatched:

$$\theta_{\text{tilt_cmd}} = \theta_{\text{tilt_gmd}} - \epsilon_{\text{pan_geom}}$$

An analogous, smaller correction is applied to the pan axis where required, and additional zoom-dependent offsets are looked up from a per-camera table to absorb the small parallax between the optical centre and the rotation axes of the gimbal at high zoom levels. After calibration, the residual pointing error is consistently below the angular field of view at maximum zoom, which is the practical requirement for keeping the target locked on screen.

Command and Protocol Layer

The system commands cameras through two complementary protocols: ONVIF for portable, vendor-independent PTZ control, and a vendor local API (LAPI) for advanced features that fall outside the ONVIF standard. The choice of protocol per command is driven by the capability required, not by the camera model.

Two capabilities required by the system are not exposed by ONVIF AbsoluteMove in a portable form: absolute focus position commanding, and dynamic on-screen-display overlays. For these the system uses the vendor LAPI, a JSON-over-HTTP interface authenticated with HTTP Basic. OSD overlays are configured by writing to the endpoint, which accepts up to eight configurable text or graphic regions per channel.

The hybrid use of ONVIF and a vendor API is a deliberate design choice. ONVIF provides portability across multiple manufacturers and enables the system to drive cameras of different makes from the same code base; the vendor API provides direct access to features that the standard either omits or implements inconsistently across vendors. The two pathways coexist on every camera that supports them.

Operator Interface and OSD Overlay

The operator interface is the layer through which test engineers interact with the system during operations. From a single console the operator can assign each camera to any of the active aircraft, fine-tune the per-camera offsets, enable or disable the position prediction and the autozoom logic, and monitor the health of the telemetry stream. The interface is intentionally kept simple, with a single control panel rather than a tabbed application, so that adjustments can be made under the time pressure of an active sortie.



Figure 7: Automated Tracking Software tracking Hurjet

A second feedback channel is provided through OSD overlays drawn directly on each video stream. Eight configurable regions display, in real time, the camera identity, the system status

flags (telemetry valid, prediction active, autozoom enabled, source available), the current target latitude, longitude, altitude and slant range, the data timestamp, and the camera location. A central crosshair is rendered from two of the regions for aiming reference. Because these overlays are produced by the camera firmware itself rather than by post-processing the video, they remain visible on every recording made from any video distribution path, including direct connections to cockpit display systems and to flight test control rooms.

Multi-Camera and Multi-Target Operation

The system supports up to eight concurrent target tracks and up to eight cameras, with arbitrary assignment between them. Multiple cameras can be assigned to the same aircraft—a common configuration during high-priority test events where redundant coverage is required—or each camera can be assigned to a different aircraft..

Internally, target updates and camera commands are dispatched in parallel using asynchronous tasks, with all camera commands issued and awaited within a single 250 ms control cycle. Because each camera operates from absolute commands rather than incremental updates, transient errors on one camera do not perturb the others, and a camera that has temporarily lost connectivity rejoins the tracking solution automatically as soon as the network path recovers.

Conclusion

This paper has described a complete telemetry-cued visual tracking system for flight test environments, built around standard ONVIF PTZ cameras and driven entirely by the live aircraft telemetry stream. The system has been deployed operationally with eight cameras and has tracked a wide spectrum of platforms, from agile jet trainers to long-endurance unmanned aircraft and helicopters in autorotation tests. All geometric computations reduce to closed-form spherical trigonometry, which removes the dependency on GPU hardware and on image-processing licenses that characterises most existing tracking solutions.

The principal engineering contribution is the recognition that the dominant error sources in a telemetry-cued PTZ system are not random but structured, and therefore can be removed with very simple regression models: a sinusoidal regression for the mechanical tilt-axis misalignment, a logistic regression for the lens distance-to-focus relationship, and a forward propagation of the aircraft position to compensate for the cumulative pipeline latency. None of these compensations require runtime

feedback from the image; they are all parameterised once during installation and remain valid for the life of the deployment, subject only to seasonal recalibration if the mechanical mounting changes.

Several limitations remain. At engagement ranges beyond a few tens of kilometres the spherical-Earth haversine formulation will need to be replaced by an ECEF-based computation that handles atmospheric refraction and ellipsoidal Earth effects more rigorously. The system is currently dependent on a healthy telemetry stream and does not yet fall back to image-based tracking when telemetry is lost; integrating a lightweight visual tracker as a backup would close this gap. A Kalman-filtered fusion of telemetry with the image-domain measurement would also tighten the residual pointing error during high-acceleration maneuvers. Finally, automatic recalibration—re-fitting the sinusoidal and sigmoid models from operational tracking data without operator intervention—would remove one of the few remaining manual procedures and is the next development priority.

Future Works

The current implementation relies on manual target designation via the ground control interface, requiring operator input to initiate tracking, and is limited to ONVIF-compatible PTZ cameras. To enable autonomous operation, improve tracking robustness, and expand hardware compatibility, future iterations will implement a three-tier hierarchical tracking architecture alongside support for legacy Pelco-D protocol cameras. The base tier will continue to use GPS-based geometric prediction via Haversine distance calculations and tilt angle compensation, providing coarse but reliable positioning even under complete visual loss. The second tier will employ a Kalman filter for motion prediction and smoothing, compensating for command latency and bridging short-duration occlusions. The top tier will integrate the CSRT (Discriminative Correlation Filter with Channel and Spatial Reliability) tracker for fine-grained visual lock, providing pixel-level accuracy under partial occlusion and scale variation. The system will dynamically degrade through the tiers: CSRT provides control when the target is visible; upon track loss, the Kalman filter maintains predicted motion; if prediction confidence drops, the system reverts to GPS-only geometric tracking. Additionally, extending protocol support beyond ONVIF to include Pelco-D will enable integration with legacy PTZ camera systems widely deployed in existing ground station infrastructure, reducing hardware replacement costs and accelerating field deployment. This

layered approach ensures continuous camera pointing even during adverse conditions, reduces operator workload during multi-aircraft scenarios, and enables fully autonomous operation without manual intervention.