

MirageMarker: A Black-Box Pose Estimation Attack for Autonomous UAV Landing

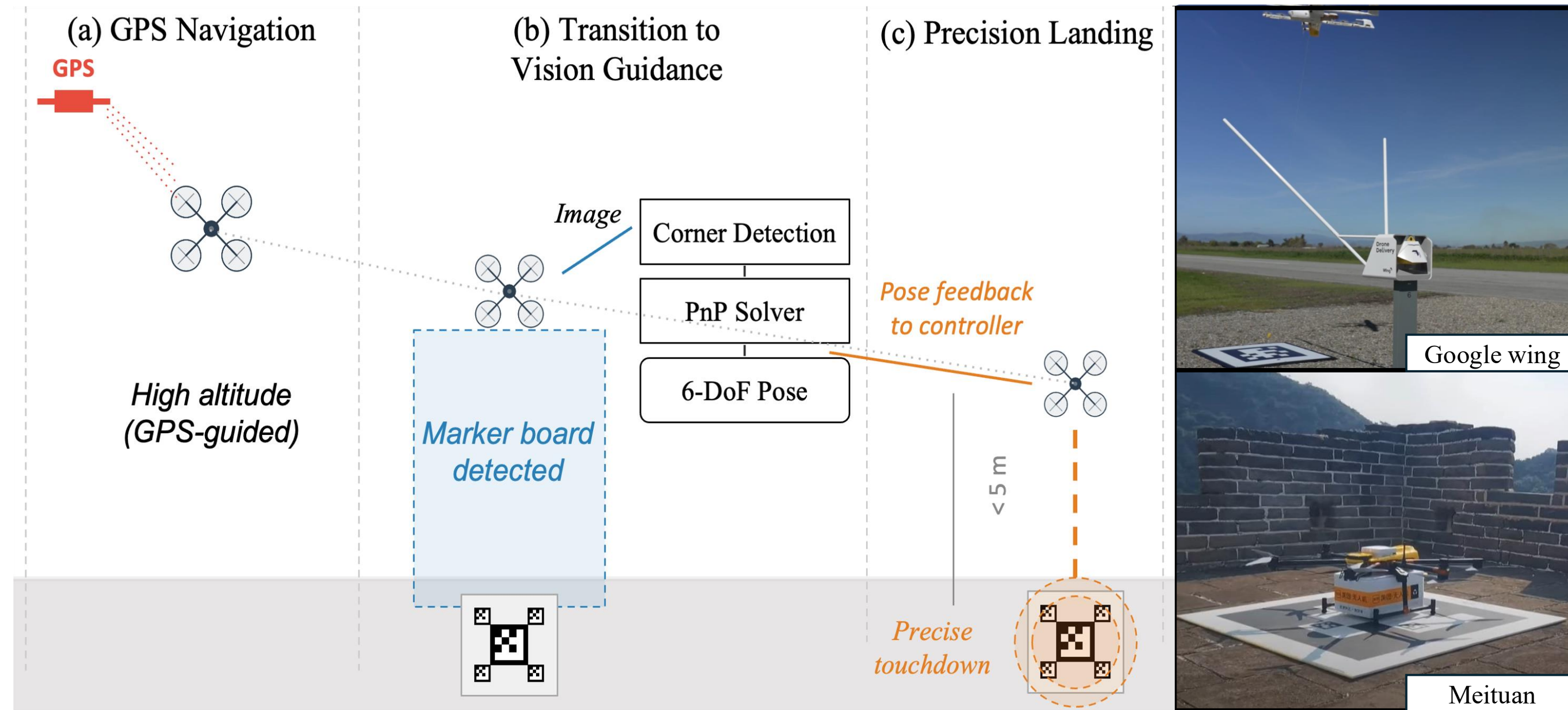
Junchi Lu, Fayzah Alshammari, Shaoyuan Xie, Xiaoqing Liang, and Qi Alfred Chen

University of California, Irvine

Motivations

ArUCo-Based Landing Systems are widely adopted in autonomous UAVs for precise localization and landing when GPS accuracy is insufficient. They provide robust, fast, and low-cost pose estimation, making them practical for applications.

- Drone Delivery: Used by companies like Google Wing and Meituan for accurate pad localization during autonomous landing [1] [2].
- Infrastructure Inspection: Deployed in industrial and urban inspection tasks requiring centimeter-level precision.



Goal & Novelty: Existing work largely focuses on marker detection robustness, while overlooking attacks on downstream fiducial marker-based pose estimation. We propose **MirageMarker** to address this gap with two key novelties:

- The first systematic study of adversarial marker-geometry manipulation against fiducial marker-based pose estimation, exploiting the fact that marker geometry is fundamental to pose estimation, with a focus on ArUco-based UAV landing.
- Formulate a novel black-box multi-objective attack for the non-differentiable ArUco pipeline, preserving detectability, physical feasibility, and stealthiness.

Threat Model

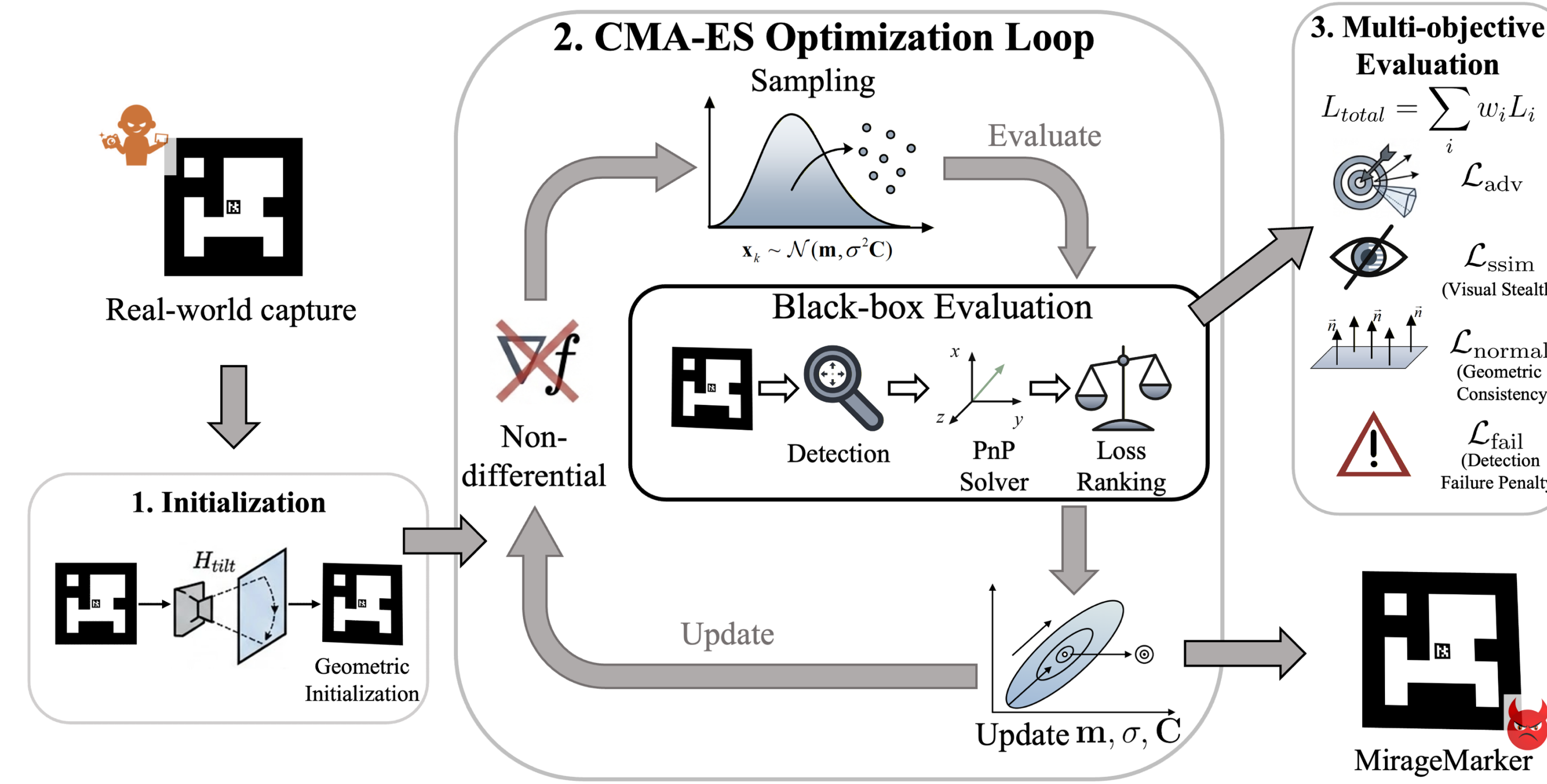
Black-box assumption:

- Temporary physical access to the landing pad enables the marker board to be measured, reproduced, and replaced or overlaid with an adversarial version. This is realistic because delivery UAV landing pads are often deployed in open and physically accessible areas [1][2]
- No assumed attacker access to the target UAV's flight-control software, onboard sensors, or controller during deployment.
- No assumed attacker knowledge of the victim beyond (1) its use of ArUco-based pose estimation; (2) its camera intrinsics or a set of possible ones including it, which can be obtained via the dataset or calibration of the same camera model after knowing or narrowing down the model from public materials or close-range inspection (e.g., during the landing) of the UAV [3][4]. MirageMarker is generated offline using black-box optimization on a general open-source PnP pipeline and relies on its transferability to the victim pipeline (shown in our evaluation).

Attack scenario:

- The attack targets the final landing phase (~2–5 m), where ArUco-based pose estimation provides precision guidance [1].

Methodology



Geometric Initialization.

Generate a view-specific adversarial marker by exploiting the sensitivity of pose estimation to small corner shifts.

Black-Box Optimization.

Use CMA-ES to optimize corner perturbations through the non-differentiable ArUco detection–PnP pipeline.

Attack Objectives.

Directional Max Attack (DMA): Push the estimated pose along the target direction $\mathbf{v}_{\text{dir}}$ by maximizing the projected horizontal pose-estimation error.

$$L_{\text{adv_max}} = -(\mathcal{P}(\mathbf{p} + \mathbf{x}) - \mathbf{T}_{\text{gt}})_h \cdot \mathbf{v}_{\text{dir}}$$

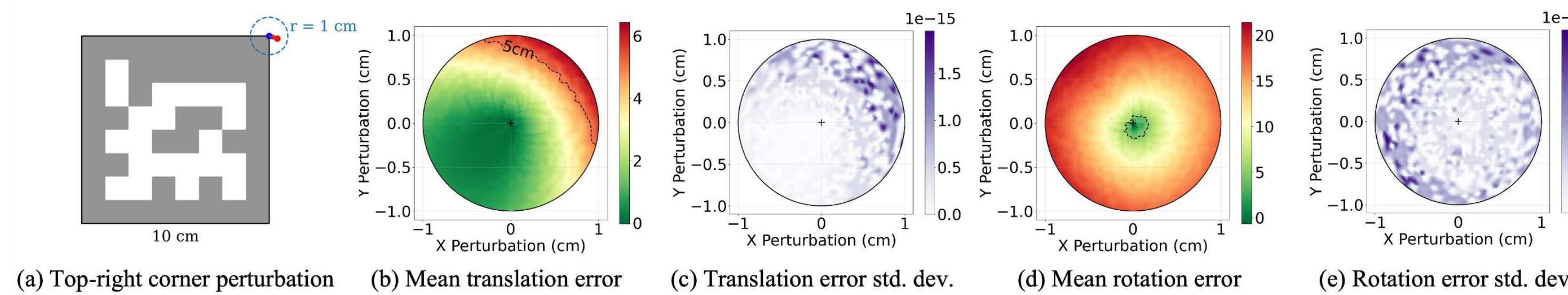
Targeted Frustum Attack (TFA): Pull the estimated horizontal translation toward zero across a direction-conditioned viewpoint frustum Φ , causing the UAV to continue along the target azimuth.

$$L_{\text{adv_tgt}} = \sum_{v \in \Phi} \tilde{w}_v \|\mathcal{P}(\mathbf{p} + \mathbf{x}, v)_h\|_2$$

Stealthiness Constraints.

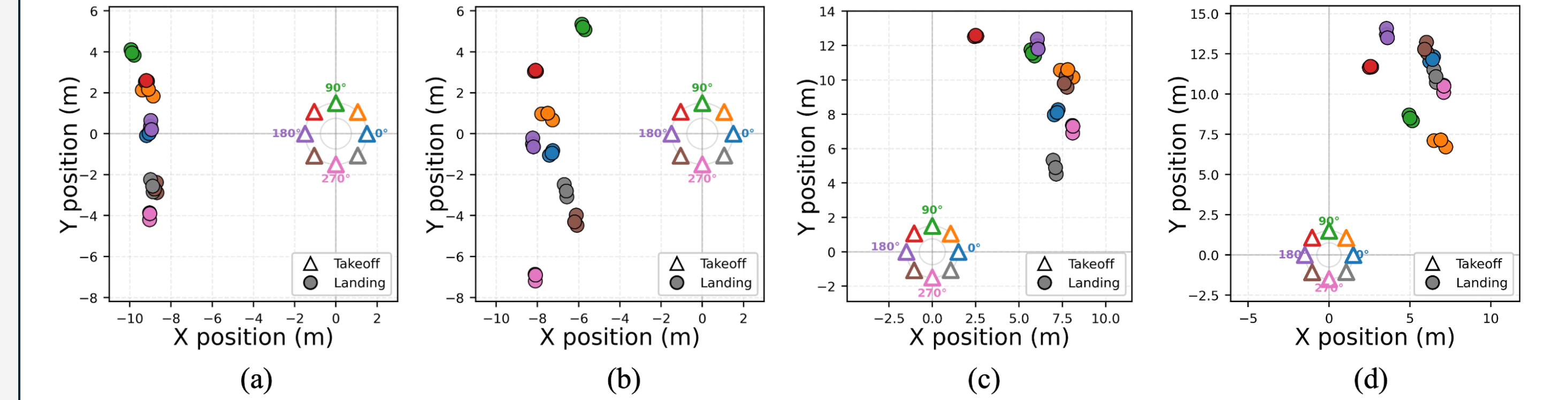
Preserve visual similarity, planar consistency, and reliable marker detectability throughout optimization.

Pose Sensitivity to Corner Perturbations

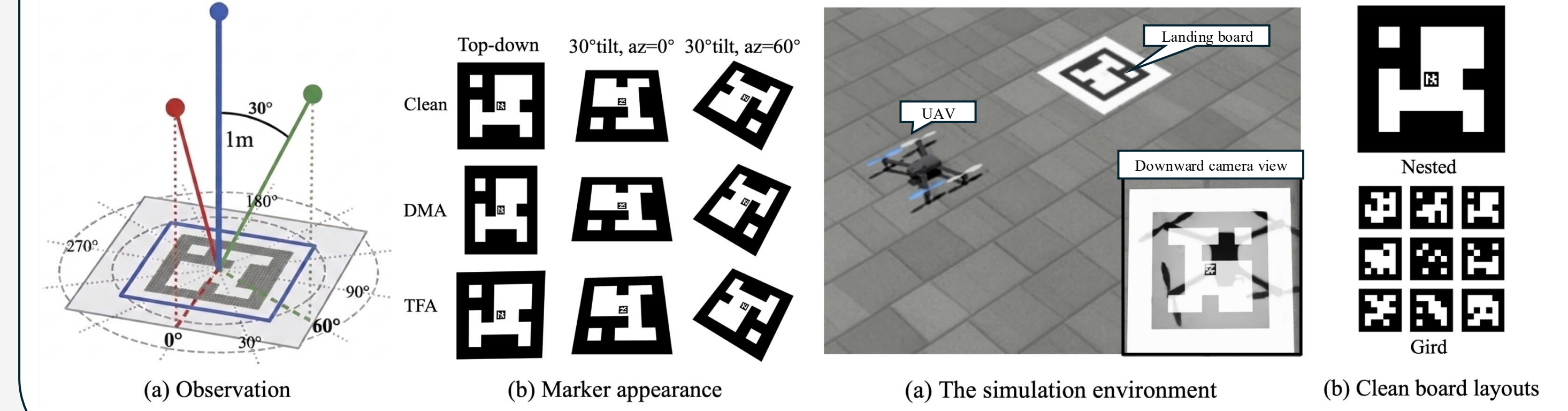


Sub-centimeter perturbations to a single marker corner can induce translation errors exceeding 5 cm and rotation errors exceeding 20°.

End-to-End UAV Landing Evaluation



Takeoff (triangles) and touchdown (circles) locations across eight azimuths: (a) nested board with DMA, (b) grid board with DMA, (c) nested board with TFA, and (d) grid board with TFA. DMA/TFA increased the touchdown error from ≤ 0.03 m to 8.12–12.37 m, causing meter-scale landing deviations. Such deviations drive the UAV outside the landing pad and could cause wrong location delivery, package loss, or unsafe landing [1][2].



Robustness to Conventional Filtering

Layout	Method	1.5 m		3 m		5 m	
		Reject.	H-err	Reject.	H-err	Reject.	H-err
Board	Clean	0/36	0	0/36	0	0/36	0
	DMA	11/36	0.49	4/36	1.19	5/36	2.05
	TFA	0/36	0.08	1/36	0.58	1/36	1.14
Nested	Clean	0/8	0	0/8	0	0/8	0
	DMA	2/8	0.30	2/8	0.95	2/8	2.09
	TFA	0/8	0.09	0/8	0.27	0/8	0.88

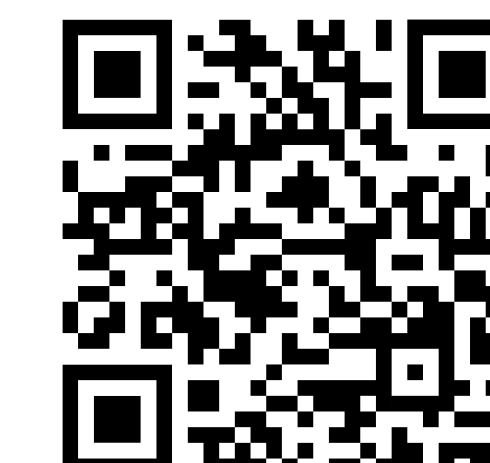
RANSAC

- RANSAC for spatial outlier rejection and KF χ^2 gate for temporal filtering.
- Most perturbed corners remain RANSAC inliers, while the KF gate passes 96.6–99.5% of attacked updates, showing that conventional filtering is insufficient.

KF χ^2 Gate

Limitation & Future Work

- Now evaluated in high-fidelity NVIDIA Isaac Sim; future work can further validate on real UAV systems to assess physical-world transferability.
- Camera models and intrinsics of commercial delivery UAVs can be difficult to obtain; the effectiveness of either (1) attack over a set of possible intrinsics with the actual one included, or (2) transfer to unknown intrinsics, remains to be evaluated.
- Non-planar marker layouts are used in indoor robot localization [5]; generalization beyond planar landing boards remains to be evaluated.



- [1] Wing, "Meet the Drones Taking Delivery to New Heights." <https://wing.com/technology>
- [2] Meituan, "A Tech-driven Retail Company." <https://www.meituan.com/en-US/about-us>
- [3] Dexta, "Meet the DE2020 — Dexta's Current Delivery Drone." <https://www.linkedin.com/posts/meet-the-de2020-dexas-current-delivery-ugcPost-7420144138253127681-3qcL/>
- [4] United UAV, "UD80 80kg Payload Transport Drone — UAV for Delivery." <https://store.uniteduav.com/products/ud80-80kg-payload-delivery-drone>
- [5] M. Kalaitzakis et al., "Fiducial Markers for Pose Estimation: Overview, Applications and Experimental Comparison," Journal of Intelligent & Robotic Systems, 2021.