

CAMERA PARAMETER SEARCH (CPS)

A Synthetic Dataset for Single-Image Camera Calibration with Ground-Truth Ininsics and Brown-Conrady Distortion

Faiz Muhammad Chaudhry¹ Jarno Ralli¹ Jerome Leudet¹ Fahad Sohrab² Farhad Pakdaman² Pierre Corbani¹ Moncef Gabbouj²

¹ AILiveSim, Helsinki, Finland

² Tampere University, Tampere, Finland

Why a new calibration dataset?

- Intrinsic and lens distortion define the 3D → 2D mapping; small errors break metric depth, SfM and visual odometry downstream.
- Learning-based single-image calibration needs dense, per-image labels, but KITTI and Cityscapes ship one calibration file per sensor, rectified images, and no explicit Brown-Conrady coefficients.
- **No public dataset provides large-scale, image-level ground truth for full intrinsics and Brown-Conrady distortion suitable for supervised learning.**

Contributions

- **CPS dataset:** ≈1.5 M images, 794 camera configurations, image-level ground truth for H-FOV, (k_1, k_2, k_3, p_1, p_2) and principal point (c_x, c_y).
- **AILiveSim generation pipeline** with constraint-based distortion sampling that guarantees geometric validity.
- **Standardised benchmark task**, evaluation protocol and baseline results as reference points for future work.

Data generation pipeline

1 Render

AILiveSim urban scenes from an ego-vehicle at 30 Hz; H-FOV 30° to 150° in 10° steps, no distortion

2 Sample

Brown-Conrady coefficients bounded to ≤ 30 px displacement at the image edge

3 Warp

Apply radial + tangential distortion and principal-point shift to the base renders

4 Clean up

Zoom & shift to crop undefined regions; store effective H-FOV, K and (c_x, c_y) in a calibration file



Fig. 1. (a) Undistorted 100° H-FOV image, (b) after random Brown-Conrady distortion, (c) after zoom-and-shift: invalid regions removed, geometry preserved.

Constraint-aware distortion sampling

Unconstrained coefficients quickly produce implausible warps, and the effect is strongly coupled with field of view: a k_1 that is harmless at 40° destroys a 150° image (right).

$$f(\sigma) = \|K d(\sigma) - K u\| - d_{\max}$$

with $d_{\max} = 30$ px

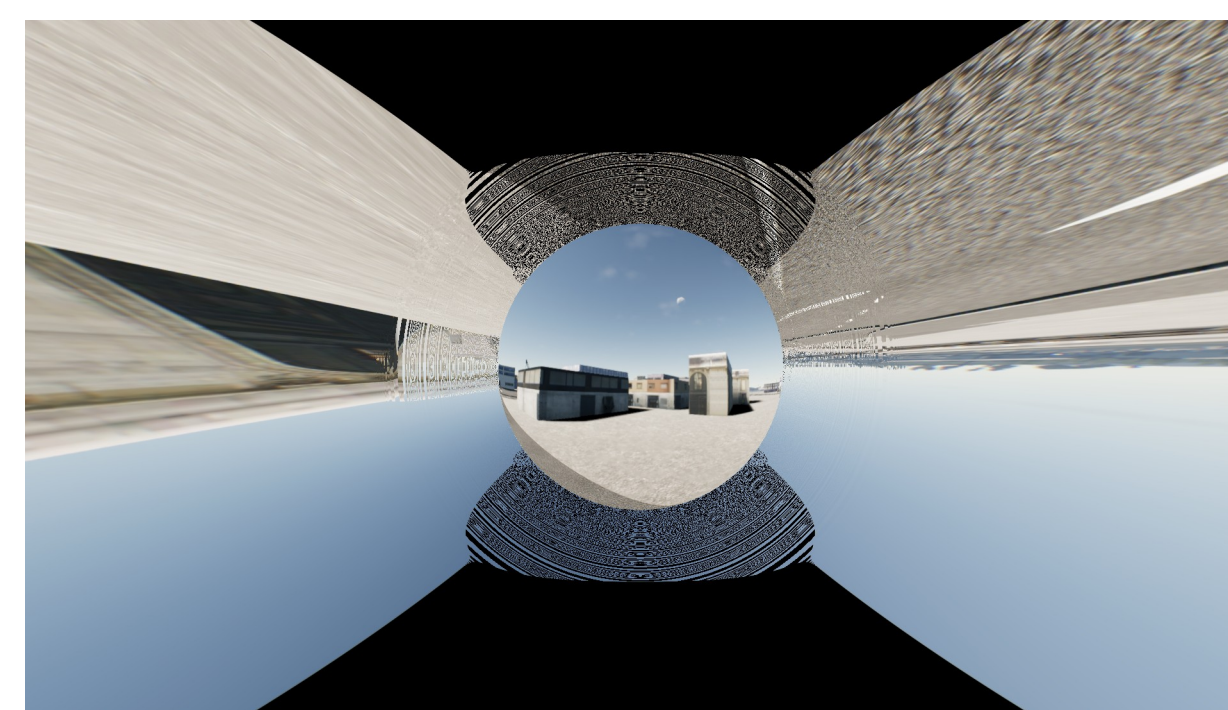


Fig. 2. $k_1 = -0.2$ at 150° H-FOV.

- For a reference point u near the image boundary (where distortion is strongest), solve $f(\sigma) = 0$ by bisection for each $\sigma \in \{k_1, k_2, k_3, p_1, p_2\}$ to get an admissible interval.
- Sample uniformly inside that interval, then move to the next coefficient; sequential sampling keeps every image physically plausible at every H-FOV.

Per-image ground truth

(H-FOV, $k_1, k_2, k_3, p_1, p_2, c_x, c_y$)

+ width, height and K, one calibration.json per configuration folder

References

- [1] Chaudhry et al., Deep-BrownConrady, IEEE T-ASE 22, 2025. [2] Leudet et al., AILiveSim, COMPSAC 2019.
- [3] Brown, Photogramm. Eng. 1971; Conrady, MNRAS 1919. [4] Geiger et al., KITTI, CVPR 2012; Liao et al., KITTI-360, TPAMI 2021.
- [5] Cordts et al., Cityscapes, CVPR 2016; Blanco et al., Málaga urban dataset, IJRR 2014. [6] Bogdan et al., DeepCalib, CVMP 2018; Tirado-Garín & Civera, AnyCalib, 2025.

CPS at a glance

1.5 M

images

30°

min horizontal FOV

150°

max horizontal FOV

2

resolutions

Resolution	Configurations	Images / folder	Total images
1920 × 1080 (16:9)	79	10,000	790,000
1392 × 512 (wide)	715	1,000	715,000
Total	794		1,505,000

Table 1. Dataset scale across the two resolution variants.



H-FOV 29.0°, $k_1 = 0.384$

H-FOV 88.2°, $k_1 = 0.003$

H-FOV 147.6°, $k_1 = 0.0006$

Fig. 3. Sample distorted CPS images across narrow, mid and ultra-wide fields of view.

Coverage of the calibration space

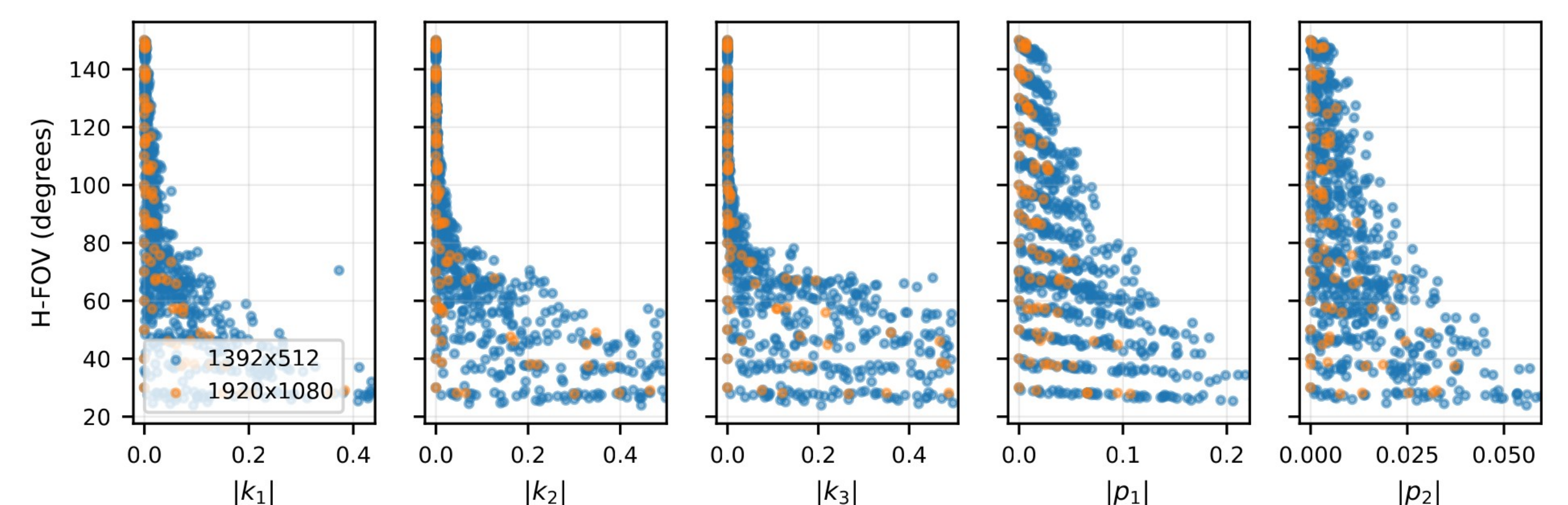
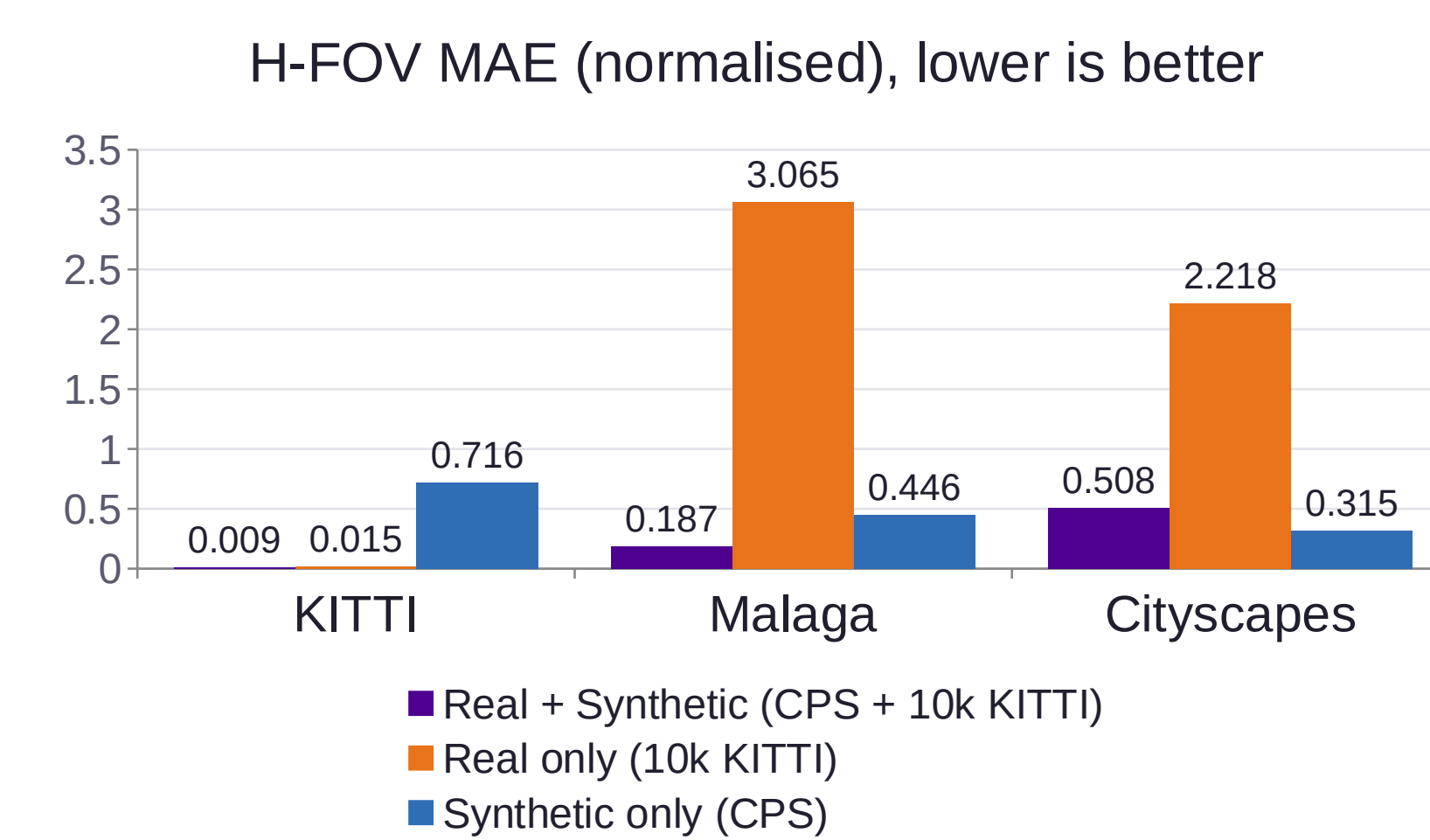


Fig. 4. |coefficient| vs. H-FOV for both resolutions. Admissible distortion ranges shrink naturally for wide-angle lenses under the 30 px displacement constraint.

Does CPS transfer to real cameras?

Task: regress the 8-parameter vector from one RGB image. DeepBrownConrady (DBC) is trained under three data regimes with identical architecture and schedule; we report per-parameter MAE (normalised space) on KITTI (in-distribution) and Malaga / Cityscapes (out-of-distribution).



MAE	Both	Real	Syn
KITTI k_1	0.008	0.004	0.374
KITTI c_x	0.002	0.005	0.009
Malaga k_1	0.150	2.969	0.340
Malaga c_x	0.017	3.458	0.009
Citysc. k_1	0.182	2.918	0.050
Citysc. c_x	0.037	2.999	0.029

Table 2. MAE excerpt (best in bold). Both = CPS + KITTI. Full table in the paper.

Main result

Training on CPS cuts out-of-distribution H-FOV error by **16× on Malaga** and **4× on Cityscapes** vs. real-only training. Best overall: **CPS + a small real set.**

Take-aways

- Real-only training over-fits to KITTI; synthetic-only is robust OOD but shows a sim-to-real gap on KITTI.
- CPS decouples scene content from camera parameters: a controlled test-bed for domain adaptation and hybrid training.
- **Dataset, calibration files and benchmark protocol are public. Scan to download.**



CPS on GitHub