Journal of Building Design and Environment
https://doi.org/10.70401/jbde.2026.0028

Supplementary information

Vision-based relative positioning in targetless environments via datumconstrained forward intersection

Yushu Yang, Gangyan Xu, Changsheng Qu, Heng Li, Haosen Chen, Lei Hou, Guomin Zhang, Wenkang Guo

Table S1. Abbreviations used in this paper.

Abbreviation	Definition
DFI	Datum-constrained forward intersection
DFI-MScale	Minimalist in-situ metric scaling in DFI
BA	Bundle adjustment
BBR	Baseline-to-range ratio (*B/Z*)
A95	95% confidence ellipse area on the site datum
GNSS	Global Navigation Satellite System
SE(3)	Special Euclidean group in 3D (rigid pose)
ChArUco	Checkerboard-ArUco hybrid calibration board
ORB	Oriented FAST and Rotated BRIEF
LoFTR	Detector-free local feature matcher
RoMa	Robust matcher (learning-based)
GT	Ground truth
LSQ	Least squares

Table S2. Symbols and subscripts used in this paper.

Symbol	Definition
A_{95}	Area of the 95% confidence ellipse on the site datum
B	Stereo baseline length between optical centres
B/Z	Baseline-to-range ratio
$\boldsymbol{C}_i$	Optical centre of camera i ($\boldsymbol{C}_i = -\mathbf{R}_i^{\mathrm{T}}\mathbf{t}_i$)
d	Signed offset of the datum plane Π: $\mathbf{n}^{\mathrm{T}} = \mathbf{x} + d = 0$
f	Focal length in pixels
$h_{\mathrm{meas},i}$	Tape-measured optical-centre height of camera i above Π
N	Number of observations or pairs (context-specific)
$\boldsymbol{P}_{bj}$	ChArUco point j on board b (board local frame)
$\mathbf{P}_{\Pi}$	Orthogonal projector onto Π ($\mathbf{I}_3 - \mathbf{n}\mathbf{n}^{\mathrm{T}}$)
$\mathbf{R}_i$	Rotation of camera i (world ← camera)
rbase	Baseline prior residual $\|\mathbf{C}1 - \mathbf{C}2\|$ - Bmeas
$r_{h,i}$	Height prior residual $\mathbf{n}^{\mathrm{T}}\mathbf{C}_i + \mathrm{d} - h_{\mathrm{meas},i}$

$\mathbf{r}_{i,j}^{reproj}$	Reprojection residual of point j in camera i
s, $\widehat{S_{lsq}}$	Metric scale correction (median or least-squares)
$\mathbf{t}_i$	Translation of camera i (world frame)
$\mathbf{T}_i^W$	Extrinsic of camera i (camera→world), $[\mathbf{R}_i \mid \mathbf{t}_i] \in \mathrm{SE}(3)$
$\mathbf{T}_b^W$	Pose of ChArUco board b in world frame
$\mathbf{u}_{ij}$	Undistorted pixel of point j in camera i
$\widetilde{\boldsymbol{u}}_i$	Homogeneous pixel direction $[u_i, v_i, 1]^{\mathrm{T}}$ (undistorted)
v_i	Unit bearing (ray) of camera i in world frame
w_i	Incidence weight $\alpha \lvert\mathbf{n}^{\mathrm{T}}\mathbf{v}_i\rvert$
$\mathbf{X}_i$	Ray-plane landing from camera i on Π
$\overline{\boldsymbol{X}}$	Incidence-weighted mean of landings
$\mathbf{X}_\Pi$	Orthogonal projection of X onto Π
$\mathbf{x}$	3D point in world frame
$\mathbf{x}_0$	Site mark (origin) on the datum plane
Z	Range (standoff) along the ray from camera to target
α	Inter-ray intersection angle at the landing point
$\chi^2_{2,0.95}$	2D chi-square quantile for 95% confidence
$\xi(\mathbf{x})$	On-datum 2D map coordinates of x
λ_i	Ray parameter for camera i: $-\frac{\mathbf{n}^{\mathrm{T}}\mathbf{C}_i + d}{\mathbf{n}^{\mathrm{T}}\mathbf{v}_i}$
$\mathbf{n}$	Unit normal of the datum plane Π
$\pi(\cdot)$	Pinhole projection operator
$\rho_{\mathrm{px}}, \rho_{\mathrm{B}}, \rho_h$	Robust losses for pixels, baseline and height priors
Σ_θ	Covariance of retained state blocks (poses, plane)
$\Sigma_{X,i}$	Landing covariance from camera i
$\Sigma_{X\Pi}$	Planar covariance of fused landing on Π
Σ_ξ	2D covariance in $(\mathbf{u}_\Pi, \mathbf{v}_\Pi)$ basis
σ_{px}	Per-axis pixel noise after undistortion
θ	Parameter/state vector used in propagation
τ	Incidence gate threshold for $\lvert\mathbf{n}^{\mathrm{T}}\mathbf{v}_i\rvert$

$\mathbf{u}_{\Pi}$, $\mathbf{v}_{\Pi}$	Orthonormal basis spanning the datum plane Π
i,b,j	Indices: camera *i*, board *b*, corner *j*; *A, B* camera labels; (*a,b*) pair index
Π, GT, est, lsq	Projected on Π; ground truth; estimate; least squares
$(\cdot)^{w}$	Quantity in world frame
$(\cdot)^{T}$, $(\cdot)^{-1}$, $(\cdot)'$	Transpose; inverse; after scale correction
$\hat{(\cdot)}$	Estimate