Contour-Anchored Transparency-Aware Ranging for Underwater Stereo

Foundation Matchers, Flat-Port Refraction, and a Negative Result for Homogenisation

Aaron Smiles

Extended Abstract · July 2026

Status: extended abstract, drafted ahead of the full write-up. Every quantitative claim traces to a logged evidence row and an experiment output; results carry three-skeptic adversarial verification with zero refutations. Manuscript in preparation; venue to be confirmed.

Problem

Transparent objects defeat stereo matching. A conventional matcher sees through the surface and assigns the object its *background* disparity, a failure characterised by Wu et al. (ICRA 2023) as a corner case in stereo. Underwater the problem compounds: a flat air–water port refracts the optical path, shifting apparent range by a scale factor that biases every uncorrected depth estimate, while turbidity and backscatter collapse detection confidence.

No published work addresses transparent-object stereo ranging underwater — the nearest neighbours handle underwater stereo without transparency, or transparency in air. The UW-TransStereo dataset (Zenodo DOI `10.5281/zenodo.16753748`; 25 stereo recordings, more than 9,000 frame pairs, more than 5,000 labelled measurements across air and four water conditions, three bottle types, CC-BY 4.0) is to our knowledge the only real-world dataset at that intersection. Prior work on this data achieved gains through preprocessing alone, with no learned-model component and no refraction model.

Approach and findings

We report three connected results over 1,378 stereo frames, evaluated against tape ground truth on held-out transparent bottles in water ($n = 249$, drawn from recordings disjoint from those used for calibration).

1. Zero-shot foundation matchers recover transparent surfaces

Three stereo foundation models applied without fine-tuning — FoundationStereo, StereoAnywhere and RAFT-Stereo — track the true range of transparent bottles underwater where classical block matching does not. Correlation against tape ground truth on raw, uncorrected depth: $r = 0.986$, 0.991 and 0.989 respectively, against $r = -0.006$ for SGBM. The premise that transparent surfaces are unmatchable does not hold for these models.

2. A matcher-independent flat-port correction makes the reading metric

Refraction through the flat port is well modelled by an affine correction $Z = n_{\text{eff}} Z_{\text{geom}} + c$, with $n_{\text{eff}} \approx 1.36$ and $c \approx -31\,\text{mm}$. Fitted on an opaque control object and transferred leave-one-environment-out across water conditions, it reduces mean absolute error from 97 mm to

12.1 mm, beating the camera SDK's own neural depth (28.5 mm). All three foundation matchers independently refit n_{eff} to 1.35–1.39, and the ratio of in-water to in-air checkerboard focal length (1.325) confirms the same port scale from calibration alone — the correction is a property of the optics, not of any one matcher.

3. Contour-anchored sampling beats the baseline, and halves the tail

Given a matcher that already reads the surface, the remaining error is in *where on the object the disparity is sampled.* Our method (M1) segments the bottle, samples the raw dense disparity across its interior, and uses the depth step at the silhouette contour to classify surface-versus-see-through reads before pooling.

Method	**MAE (mm)**	**p90 (mm)**	**600–900 mm band**
StereoAnywhere centre-pixel (baseline)	15.0	27.6	29.7
Raw interior median (pooling only)	11.1	18.4	26.0
TA-Stereo homogenisation	12.5	16.1	28.0
Contour depth-step (ours)	**9.7**	**15.3**	**24.7**

M1 beats the baseline on 80.7% of frames, with placement-level Wilcoxon $p = 1\text{e}{-4}$. An ablation ladder separates the two mechanisms: robust interior pooling carries 74% of the gain, the depth-step surface classifier the remaining 26%. The improvement is attributable to the estimator rather than to auxiliary priors — zeroing the known-dimension half-depth offset still yields 10.0 mm MAE, and the refitted n_{eff} (1.356) is unchanged from the baseline's (1.347), so it is not a disguised refraction refit.

4. A negative result, disclosed

We reimplemented TA-Stereo homogenisation — masking the transparent region and filling it with boundary statistics so the matcher treats it as opaque — and it does not help. At 12.5 mm MAE it is *worse* than simply pooling the raw interior (11.1 mm). Adversarial verification corrected our own first-pass explanation: homogenisation does not destroy the signal (the filled region still reads the surface, 10.5 mm on the raw axis) but reads systematically nearer, on 89% of frames with median −5.4 mm, carrying a −11.7 mm bias through the ranging protocol.

The lesson generalises. TA-Stereo's premise is to opacify *because* the matcher fails on transparent surfaces, and that premise is void once the matcher already reads them. The correct move is to sample the foundation disparity better, not to overwrite it.

Contributions

1. First demonstration that zero-shot stereo foundation models recover transparent-object surface depth underwater.

2. An opaque-control-calibrated flat-port refraction correction that transfers across water conditions and is independent of the matcher.

3. A contour-anchored depth-step ranging method that reduces MAE from 15.0 to 9.7 mm and halves the p90 error tail, with an ablation separating robust pooling from surface classification.

4. A disclosed negative result for transparency homogenisation applied to foundation matchers, with a verified mechanism.

Limitations

Results are in-distribution over a 200–800 mm working range on a single camera and port geometry. The known-dimension half-depth is a per-bottle size prior derived from the data rather than a specification, though it is shown to be non-load-bearing. One water condition (freshwater with the legacy enhancement preprocessing) is the single environment where the method ties with rather than beats the baseline, and the 600–900 mm band remains the weakest for every method tested.

Extends the UW-TransStereo dataset (Zenodo DOI `10.5281/zenodo.16753748`), funded by UKRI grant 2601988. Author list pending confirmation of co-authors.