Manifold-Guided Stereo-Monocular Refinement for Endoscopic Stereo Disparity Estimation

Jiewen Yang[1] Yi Zhong[2] Bin Pu[3*] Xingbo Dong[2] Ming Zheng[5] Qika Lin[4] Kai He[4] Xuechen Zhang[1*]

[1] The Hong Kong University of Science and Technology [2] Anhui University [3] Hunan University
[4] National University of Singapore [5] Anhui Normal University

Correct with stereo geometry. Fuse where stereo is uncertain.

A reliability-gated latent-flow framework for endoscopic disparity.

Scan me!

Project Page: https://github.com/Jiewen-Yang/Manifold-Stereo

SCARED	SERV-CT	HAMLYN
IN-DOMAIN	ZERO-SHOT	ZERO-SHOT
3.05 px EPE ↓	**1.70** px EPE ↓	**4.24** px EPE ↓
10.0% lower vs best baseline	35.4% lower vs best baseline	4.5% lower vs best baseline

01 Why local reliability?

Reliable matches and ambiguous tissue coexist within the same stereo pair.

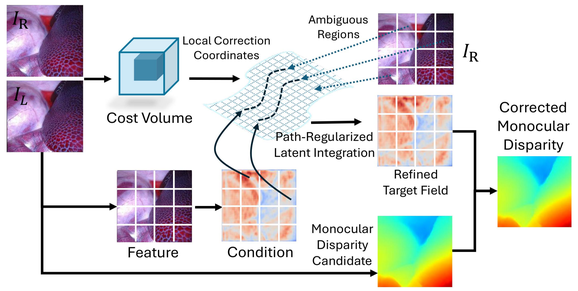


Fig. 1 · Local, stereo-conditioned correction

04 Why each part matters

Controlled ablations on SCARED · EPE ↓

Setting	px	Δ px
Full model	**3.05**	—
No correction space	3.27	+0.22
Uniform mixing	3.30	+0.25
No geometric gate	3.45	+0.40
One-step flow	3.55	+0.50
No path-energy loss	3.84	+0.79
No scale alignment	4.32	+1.27

Multi-step flow, shared parameters.
The one-step control uses the same encoder, vector field and decoder.

Tables 4–6 · Δ measured relative to the full model

02 Two coupled innovations

A - CORRECT THE CANDIDATE

Stereo-conditioned Latent Flow

Cost-volume geometry anchors local monocular corrections.

B - CONTROL ITS CONTRIBUTION

Geometric-consistency Gating

Left–right feature consistency gates scale-aligned monocular context.

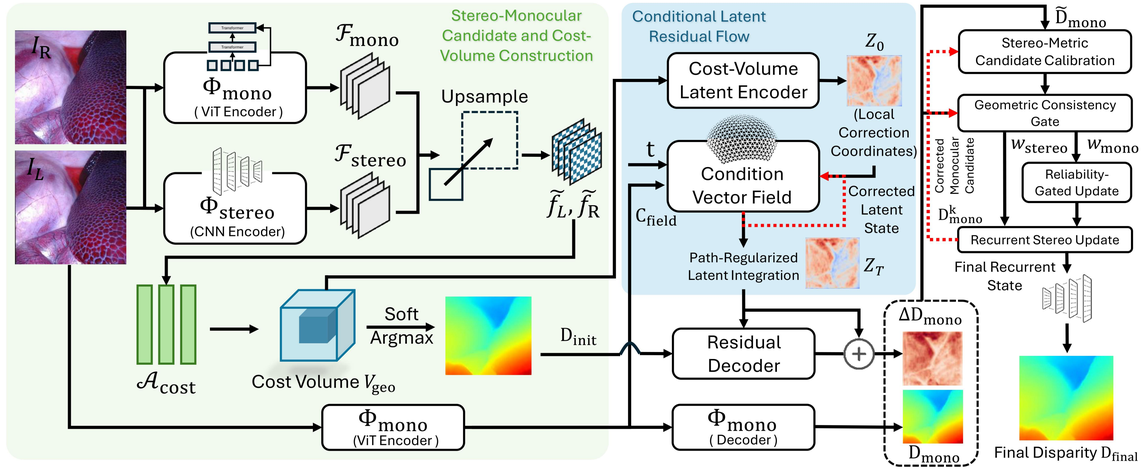
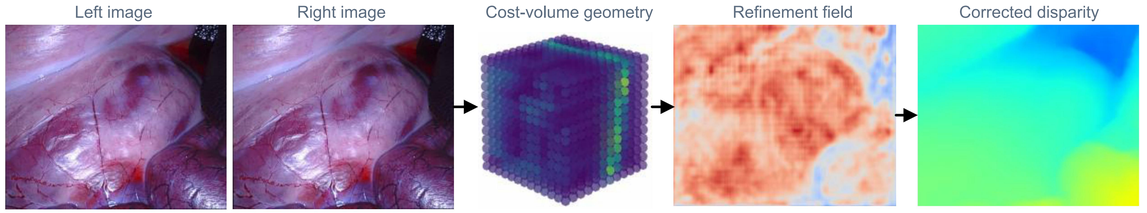


Fig. 2 · Correct the candidate → align to stereo scale → gate recurrent updates

Local correction in action

Fig. 5 · example 1



03 Accuracy that transfers

One SCARED-selected checkpoint.
No target-dataset fine-tuning.

EPE (px) ↓	SCARED	SERV-CT	Hamlyn
RAFT-Stereo	4.14	3.24	4.71
Monster	3.51	4.68	7.57
StereoMamba	3.39	2.63	4.98
RRESM	4.02	3.11	4.44
Ours	**3.05**	**1.70**	**4.24**

Selected baselines · full comparisons in Tables 1–3

Best EPE on all three datasets.
SCARED depth MAE: 2.53 mm

Zero-shot visual comparison

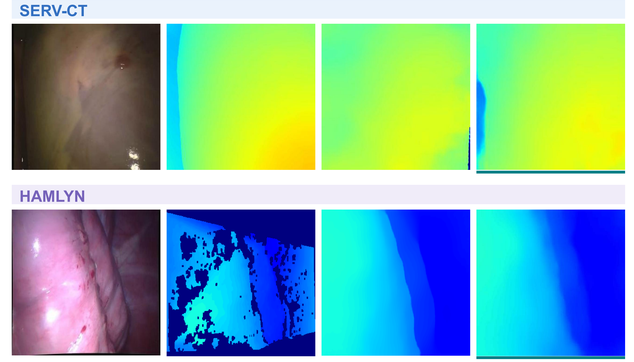


Fig. 4 · Original paper panels, selected first examples

Best Hamlyn Bad3: 27.96%.
Other Hamlyn outlier rates are competitive.

CONTACT jyangcu@connect.ust.hk

Relative EPE gains use the strongest baseline in each full table.