Variational Simultaneous Stereo Matching and Defogging in Low Visibility

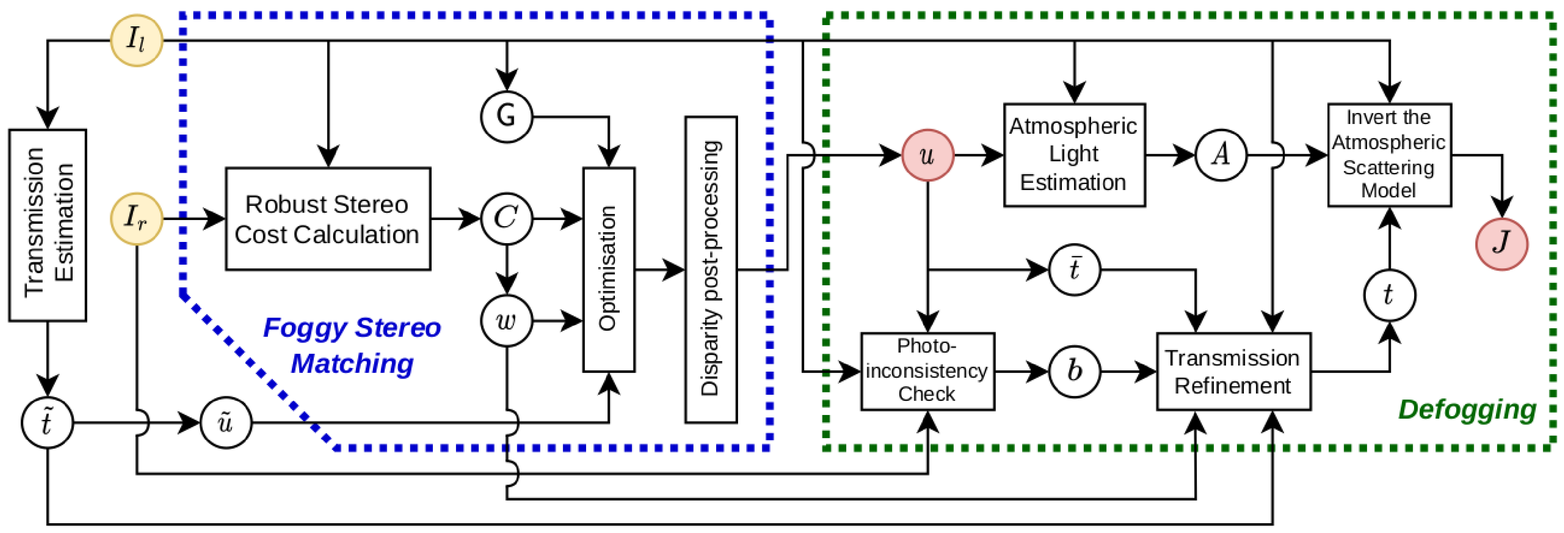
Ted (Yining) Ding[1] Andrew Wallace[1] Sen Wang[2]

[1]Heriot-Watt University [2]Imperial College London

Introduction

- PROBLEM: We seek to estimate *a dense disparity map* and *a defogged image* from *a rectified stereo pair of foggy images*.
- RESEARCH GAP: Existing stereo matching algorithms are predominantly developed under the assumption of *clear* scenes. Meanwhile, most of the literature on defogging addresses *single* images only. There is very *little* work that tackles the two problems *simultaneously*.
- MOTIVATION: These two tasks are closely linked by scene *depth*, which can be inferred from *disparity* and *scattering*, respectively.
- SOLUTION: We propose a novel algorithm which *effectively exploits this underlying connection* and *improves both results*. It is based on *variational continuous optimisation* and *does not require large scale training data*, the acquisition of which is not always possible outdoors.

Method



Input:

- I_l Left foggy image
- I_r Right foggy image

Output:

- u Left dense disparity map
- J Left defogged image

$$I_l, I_r, J \in \mathbb{R}^{H\times W\times 3}$$

$$\tilde{t}, \bar{t}, t \in [0,1]^{H\times W}$$

$$A \in \mathbb{R}^3$$

$$\tilde{u}, u, w, b \in [0,1]^{H\times W}$$

$$\mathsf{G} : \mathbb{R}^{H\times W\times 2} \rightarrow \mathbb{R}^{H\times W\times 2}$$

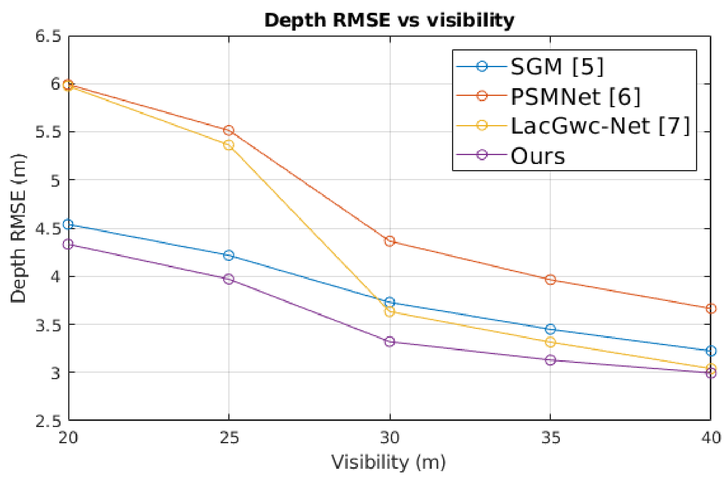
$$C \in \mathbb{R}^{H\times W\times |\Gamma|}$$

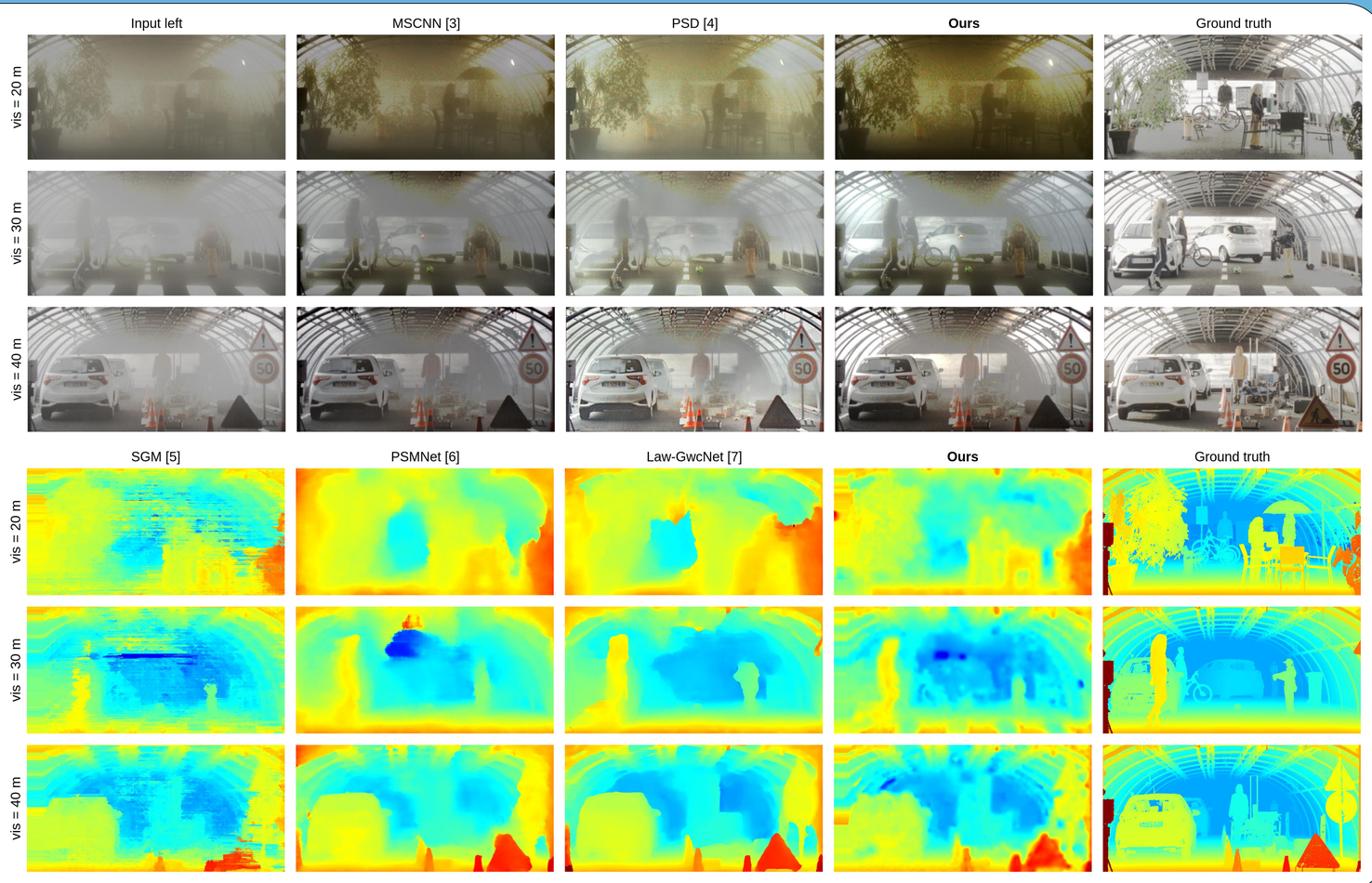
$$\Gamma = \{0, 1, \cdots, \gamma_{\max}\}$$

- Our two-stage system consists of a *Foggy Stereo Matching* module and a *Defogging* module.
- The Foggy Stereo Matching module has the following features which get seamlessly incorporated into our disparity optimisation process.
  - An anisotropic weighting scheme to allow for *non-uniform penalty parameters at different pixel locations*.
  - A customised regularisation term which effectively injects disparity cues from scattering by *encouraging gradient alignment*.
- Within the Defogging module we leverage a *photo-inconsistency check* and perform a *transmission refinement* to enable the recovered depth information to assist defogging.

Results

- We evaluate our method qualitatively and quantitatively on *both synthetic* data (VKITTI2 [1], shown in paper) and *real* data (PAD [2], shown in paper and here).
- Our algorithm *outperforms* comparative methods in all metrics on depth estimation, and produces *visually more appealing* defogged images, especially *in extremely low visibilities*.





Future Work

- Use motion information embedded in consecutive frames and incorporate more matching constraints to improve depth estimation results
- Adopt a more sophisticated fog model (*e.g.* blurring and fog inhomogeneity) to better recover intensity images

References

[1] Cabon *et al.*, "Virtual KITTI 2", arXiv preprint arXiv:2001.10773, 2020.
[2] Gruber *et al.*, "Pixel-accurate depth evaluation in realistic driving scenarios," 3DV'19
[3] Ren *et al.*, "Single image dehazing via multi-scale convolutional neural network", ECCV '16
[4] Chen *et al.*, "PSD: Principled synthetic-to-real dehazing guided by physical priors", CVPR'21
[5] Hirschmuller, "Stereo processing by semiglobal matching and mutual information", PAMI'07
[6] Chang and Chen, "Pyramid stereo matching network", CVPR'18
[7] Liu *et al.*, "Local similarity pattern and cost self-reassembling for deep stereo matching networks", AAAI'22

Acknowledgements

We thank João F. C. Mota for helpful discussions. This work is supported by EPSRC funding for the Centre for Doctoral Training in Robotics and Autonomous Systems EP/S023208/1.

Contact details

Email: yd2007@hw.ac.uk