OmniFlowNet: a Perspective Neural Network Adaptation for Optical Flow Estimation in Omnidirectional Images

Charles-Olivier Artizzu[1], Haozhou Zhang[2], Guillaume Allibert[1] and Cédric Demonceaux[2]

[1]Université Côte d'Azur, CNRS, I3S, France.

[2]VIBOT ERL CNRS 6000, ImViA, Université Bourgogne Franche-Comté, France.

Context

Goal: **Optical flow estimation in spherical images using deep learning approaches.**

Issues:

- ▶ No omnidirectional annotated dataset;
- ▶ Strong images distortions due to projection.

Solutions: **Adaptation** of perspective networks taking into account distortions by modifying the convolution.

Local perspective projection on the sphere

Adapting the perspective CNN kernels to equirectangular: (Fernandez et al. 2020)

- ▶ Spherical coordinates of **the center of the kernel**:

$$\phi_{00} = \left(u_{00} - \frac{W}{2}\right)\frac{2\pi}{W}; \quad \theta_{00} = -\left(v_{00} - \frac{H}{2}\right)\frac{\pi}{H}, \tag{1}$$

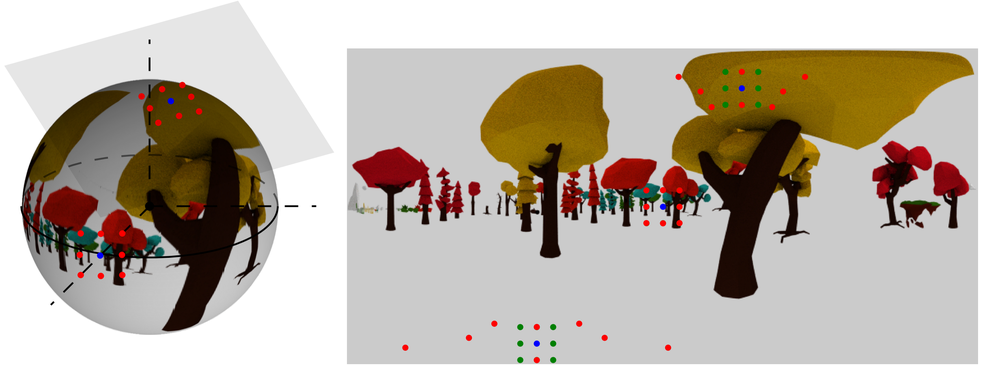
- ▶ For a kernel size r with local points (i, j) in $\left[-\frac{r-1}{2}, \frac{r-1}{2}\right]$, the coordinates of these points on the **local perspective projection** are :

$$p_{spher,ij} = \begin{bmatrix} x_{ij} \\ y_{ij} \\ z_{ij} \end{bmatrix} = R_y\left(\phi_{00}\right) R_x\left(\theta_{00}\right) \frac{\hat{p}_{spher,ij}}{|\hat{p}_{spher,ij}|}, \tag{2}$$

where $R_a(\beta)$ rotation matrix of β around the a axis.

- ▶ Finally the **back projection on the 2D equirectangular image** give these coordinates for the kernel points:

$$u_{ij} = \frac{W}{2\pi}\left(\arctan\left(\frac{x_{ij}}{z_{ij}}\right) + \pi\right); \quad v_{ij} = -\frac{H}{\pi}\left(\arcsin\left(y_{ij}\right) - \frac{\pi}{2}\right). \tag{3}$$



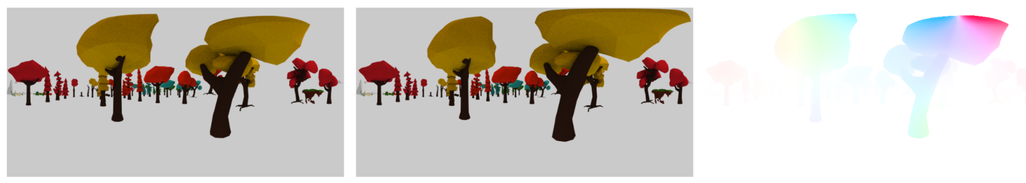
Blue: center of kernel. Green: Perspective kernel. Red: Spherical kernel.

Proposed Solution

Implemented on **LiteFlowNet2** (Hui et al. 2020).
Weights from **LiteFlowNet2** authors on Sintel training Butler et al. 2012.

Validation on virtual datasets built on Blender

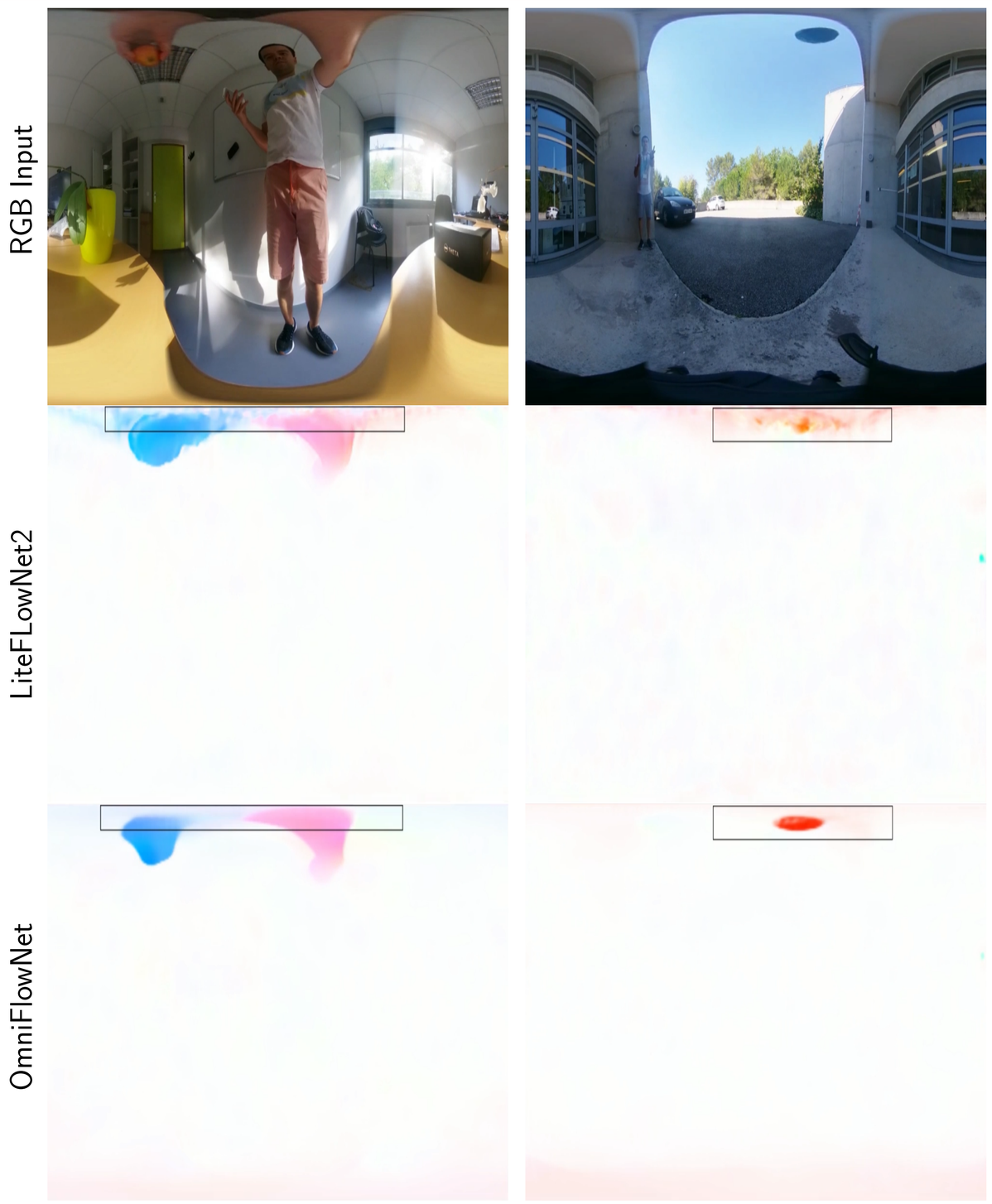
Ground truth optical flow extracted using Blender *Vector Pass* (Ranjan et al. 2020).



Comparison metrics used (Baker et al. 2007):
Angular Error (in degrees) and **Endpoint Error** (in pixels):

Average on 1200 frames	Cartoon Tree		Forest		Low Poly Model	
	AE	EE	AE	EE	AE	EE
LiteFlowNet2	63.07	5.60	58.11	10.61	58.03	7.66
OmniFlowNet	**54.74**	**4.49**	**55.05**	**9.72**	**55.21**	**7.23**

Validation on real videos



http://www.i3s.unice.fr/~allibert/Videos/icpr20_video.mp4.

Conclusion

OmniFlowNet:

- ▶ perspective CNN adapted to equirectangular images;
- ▶ plugin transferable on any CAFFE CNN;
- ▶ no extra training needed;
- ▶ no slowdown in the CNN execution;
- ▶ proven performances on real and virtual datasets.

References

Baker, S. et al. (2007). "A Database and Evaluation Methodology for Optical Flow". *International Conference on Computer Vision* 92. DOI: 10.1007/s11263-010-0390-2.

Butler, D. J. et al. (2012). "A naturalistic open source movie for optical flow evaluation". *European Conf. on Computer Vision (ECCV)*, pp. 611–625.

Fernandez, Clara et al. (2020). "Corners for Layout: End-to-End Layout Recovery from 360 Images". *IEEE Robotics and Automation Letters* PP, pp. 1255–1262. DOI: 10.1109/LRA.2020.2967274.

Hui, Tak-Wai et al. (2020). "A Lightweight Optical Flow CNN - Revisiting Data Fidelity and Regularization". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. PP, pp. 1–1. DOI: 10.1109/TPAMI.2020.2976928.

Ranjan, Anurag et al. (2020). "Learning Multi-Human Optical Flow". *International Journal of Computer Vision*. DOI: 10.1007/s11263-019-01279-w.

artizzu@i3s.unice.fr