OmniFlowNet: a Perspective Neural Network Adaptation for Optical Flow Estimation in Omnidirectional Images
Submission 833

Charles-Olivier Artizzu\textsuperscript{1}, Haozhou Zhang\textsuperscript{2}, Guillaume Allibert\textsuperscript{1} and Cédric Demonceaux\textsuperscript{2}

\textsuperscript{1}Université Côte d’Azur, CNRS, I3S, France.

\textsuperscript{2}VIBOT ERL CNRS 6000, ImViA, Université Bourgogne Franche-Comté, France.

25th International Conference on Pattern Recognition
10-15 January 2021, Milan, Italy
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.

Equirectangular image of an outdoor car scene. Polar regions are highly distorted
**Optical Flow**

**Optical flow**: estimation of the apparent motion between two images.

**Methods of computation:**

<table>
<thead>
<tr>
<th>Year</th>
<th>Method Type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>Phase-Based methods:</td>
<td>(Radgui et al. 2011)</td>
</tr>
<tr>
<td>2015</td>
<td>Supervised Deep Learning: FlowNet</td>
<td>(Dosovitskiy et al. 2015)</td>
</tr>
<tr>
<td>2020</td>
<td>LiteFlowNet2</td>
<td>(Hui, Tang, and Loy 2020)</td>
</tr>
</tbody>
</table>

Optical Flow from Sintel Dataset (Butler et al. 2012)
Adaptation to spherical

Solutions to adapt to spherical distortions:

- **Build an Omnidirectional annotated dataset**
- **Change the entire features map:**
  Fast Fourier Transformation (Cohen et al. 2018),
  Polyhedra (Lee et al. 2018)
- **Adapt the perspective CNN kernels to equirectangular:**
  Linear Adaptation (Su and Grauman 2017),
  Spherical Adaptation (Fernandez et al. 2020)
Local perspective projection of CNN kernels on the sphere:


Proposed solution available on any CAFFE (Jia et al. 2014) based CNN.
Proposed solution

Implemented on **LiteFlowNet2** (Hui, Tang, and Loy 2020). Weights from **LiteFlowNet2** authors (training MPI Sintel).

1. **Training**

   *Pair of perspective images*

   **LiteFlowNet2**

   *Estimated optical flow*

   *Perspective ground truth*

2. **Testing**

   *Pair of equirectangular images*

   **OmniFlowNet**

   *Estimated optical flow*

   *Equirectangular convolution*

   **Architecture & Weights**
Validation on virtual datasets

Virtual scenes built on **Blender**: 3 scenes, 4 camera orientations. **Ground truth optical flow** extracted using **Vector Pass** (Ranjan et al. 2020).

![Virtual scenes](image)

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

\[
AE = \cos^{-1}\left(\frac{1 + u \cdot u_{gt} + v \cdot v_{gt}}{\sqrt{1 + u^2 + v^2} \sqrt{1 + u_{gt}^2 + v_{gt}^2}}\right)
\]

\[
EE = \frac{1}{N} \sum \sqrt{(u_{gt} - u)^2 + (v_{gt} - v)^2}
\]

with \((u, v)\) the estimated flow and \((u_{gt}, v_{gt})\) the ground truth.

<table>
<thead>
<tr>
<th>Average on 1200 frames</th>
<th>Cartoon Tree</th>
<th>Forest</th>
<th>Low Poly Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AE</td>
<td>EE</td>
<td></td>
</tr>
<tr>
<td>LiteFlowNet2</td>
<td>63.07</td>
<td>5.60</td>
<td></td>
</tr>
<tr>
<td>OmniFlowNet</td>
<td>54.74</td>
<td>4.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Validation on real videos

Real scenes were filmed using a Theta Ricoh Z1:
- Apple 1 and Ball 3: moving indoor scene / fixed camera;
- Ball 1 and Ball 2: moving outdoor scene / fixed camera;
- Car 1 and Car 2: fixed outdoor scene / moving camera.

Same performances in the equatorial region.

Better performances of OmniFlowNet in the polar regions.

Complete video available on:
http://www.i3s.unice.fr/~allibert/Videos/icpr20_video.mp4.
Validation on real videos

RGB Input

LiteFlowNet2

OmniFlowNet
Conclusion

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


Thank you for your attention.