Optical Flow Requires Multiple Strategies (but only one network)

Tal Schuster, Lior Wolf and David Gadot

IEEE 2017 Conference on Computer Vision and Pattern Recognition

CVPR July 21-26 HONOLULU 2017

Main Contributions

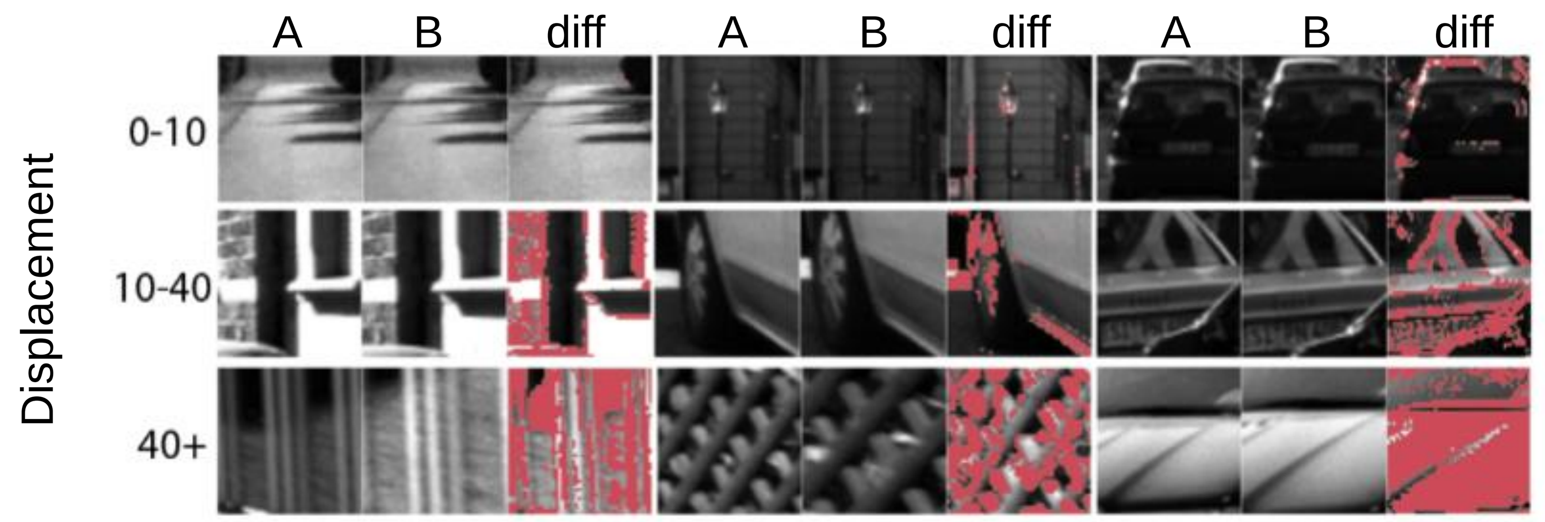
- We analyze **the need for multiple strategies in optical flow**
- We propose a novel, psychologically inspired way to **train a network to address multiple scenarios at once**
- We show how, in optical flow, our proposed new scheme translates to a simple, unexpected, heuristic
- We improve the PatchBatch pipeline
- **State of the art results** are demonstrated on the KITTI 2012 and KITTI 2015 benchmarks*

Optical Flow as a Multifaceted Problem

- Generating **descriptors** for **small and large displacements**
- **Distractor** – a pixel in the destined image that its descriptor is closer than the ground truth to the descriptor of the original image
- Distractors amount per displacement range for models trained on all the data, or trained only on patches with displacement smaller\ larger than 30 pixels:

	0 – 5	5 – 10	10 – 20	20 – 30	30 – 45	45 – 60	60 – 90	90 – ∞
All	**2.32**	7.32	5.32	9.38	25.21	50.43	67.32	216.39
< 30	2.46	**6.91**	**5.25**	**8.57**	26.39	51.76	65.15	209.40
> 30	3.03	9.07	5.64	10.29	**24.74**	**46.81**	**56.69**	**199.61**

- The **need of different features** for **different patches**:



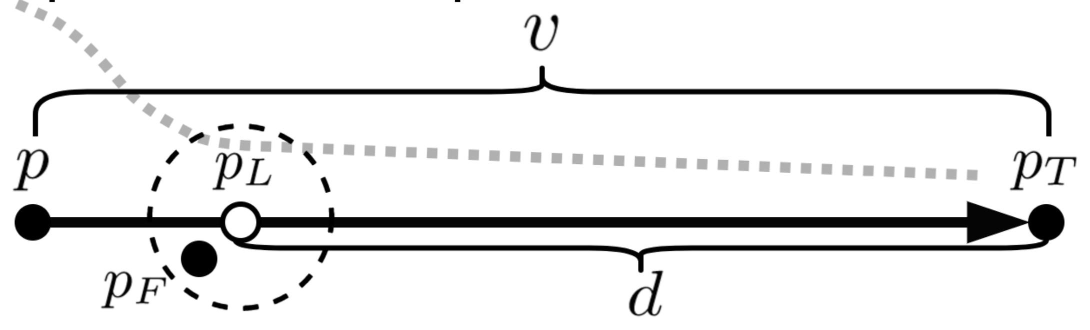
- Descriptors for large displacements are harder generate
- Known gradual methods fail due to the need of different strategies

Interleaving Learning

- **Balance the difficulty** of samples from all categories for each batch by using easier negative samples for harder positive ones

$X \sim \log\mathcal{N}(\mu,\sigma), \quad d = v(1-X)$

$P(X=x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)}$



p - pixel from Image A, p_T - positive match, p_F - negative match

Training with Triplets

$L = \sum \max(0, m + D_{pos} - D_{neg}) + \lambda$ SD

Image A

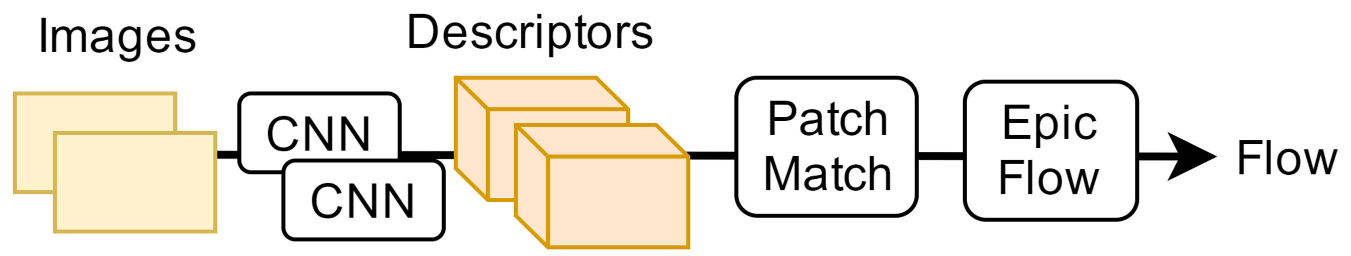
Image B Baseline

Image B Interleaving

- Positive sample

- Negative sample

The PatchBatch Pipeline



Self-Paced Curriculum Interleaving Learning (SPCI)

- Applying the **interleaving learning method** allows to gradually increase the overall difficulty

$$d_i = v(1 - X - R_i)$$

$$R_i = \underbrace{\frac{i}{m}}_{curriculum} \cdot \underbrace{\max\left(0,1-\frac{l_{i-1}}{l_{init}}\right)}_{self-paced}$$

l_i - validation lose on epoch i

m – total epoch amount

Results

- Distractors amount comparison of known gradual methods and our interleaving learning method (using Hinge loss with SD regularization component):

	0 – 5	5 – 10	10 – 20	20 – 30	30 – 45	45 – 60	60 – 90	90 – ∞	Avg.
Baseline	2.32	7.32	5.32	9.38	25.21	50.43	67.32	216.39	20.51
Neg-Mining	3.06	6.19	5.41	10.52	26.88	51.33	70.29	210.34	20.96
Curriculum	2.83	8.66	5.25	10.35	23.62	45.82	63.69	197.82	19.70
Self-Paced	2.88	9.35	6.84	13.74	34.09	57.46	80.80	198.97	23.93
Interleaving	1.41	5.57	**3.07**	**6.31**	15.60	28.52	43.46	**127.65**	12.61
SPCI	**1.40**	**5.04**	3.46	6.56	**15.11**	**27.13**	**42.72**	130.17	**12.50**

- KITTI 12'

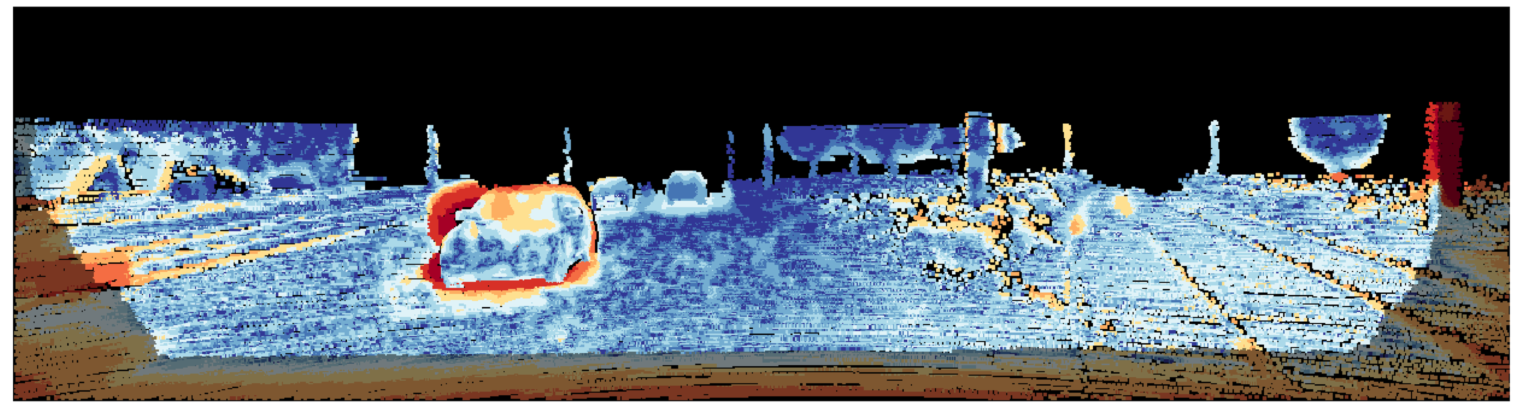
Method	Out-NOC
Ours	4.65%
Baseline	5.44%

KITTI15'

Method	Fl-bg	Fl-fg	Fl-all
Ours	17.25%	24.52%	18.46%
Baseline	19.98%	30.24%	21.69%

MPI - Sintel

Method	EPE	s0-10	s40+
Ours	6.22	0.91	39.91
Baseline	6.78	0.72	45.86



* At the time of writing