EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow

Jerome Revaud[a], Philippe Weinzaepfel[a], Zaid Harchaoui[a,b], Cordelia Schmid[a]
[a] LEAR team, Inria Grenoble Rhone-Alpes, Laboratoire Jean Kuntzmann, CNRS, Univ. Grenoble Alpes, France [b] NYU

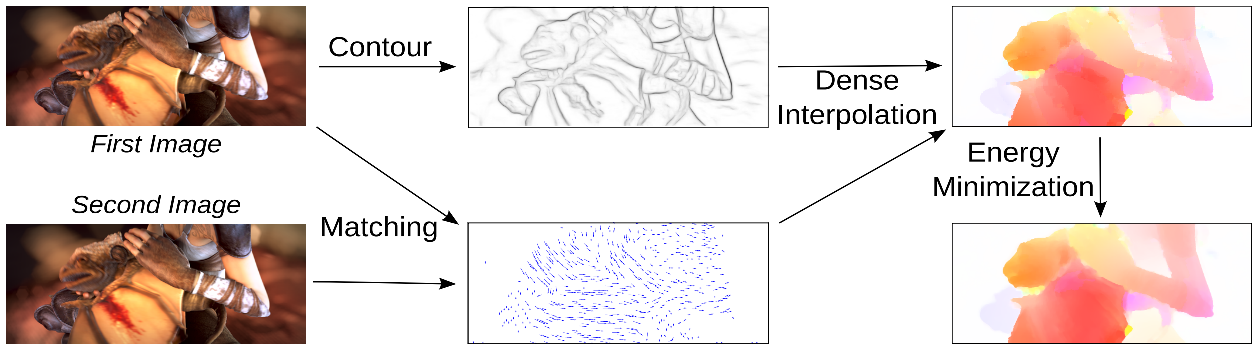

Figure 1: Given two images, we compute matches using DeepMatching [6] and the edges of the first image using SED [3]. We combine these two cues to densely interpolate matches and obtain a dense correspondence field. This is used as initialization of a one-level energy minimization framework.

Most state-of-the-art optical flow approaches are built upon an energy minimization framework, often solved using efficient coarse-to-fine algorithms [1, 5]. A major drawback of coarse-to-fine schemes is error-propagation, *i.e.*, errors at coarser levels, where different motion layers can overlap, can propagate across scales. Even if coarse-to-fine techniques work well in most cases, we are not aware of a theoretical guarantee or proof of convergence.

Instead, we propose to simply interpolate a sparse set of matches in a dense manner to initialize the optical flow estimation, see Figure 1. This novel procedure enables us to leverage recent advances in matching algorithms, which can now output quasi-dense correspondence fields [6]. In the same spirit as [4], we perform a sparse-to-dense interpolation by fitting a local affine model at each pixel based on nearby matches. Nevertheless, a major issue arises for the preservation of motion boundaries. We make the following observation: *motion boundaries often tend to appear at image edges*, see Figure 2. Consequently, we propose to exchange the Euclidean distance with a better, *i.e.*, edge-aware, distance and show that it offers a natural way to handle motion discontinuities. Moreover, we show how an approximation of the edge-aware distance allows to fit only one affine model per input match (instead of one per pixel). This leads to an important speed-up of the interpolation scheme without loss in performance.

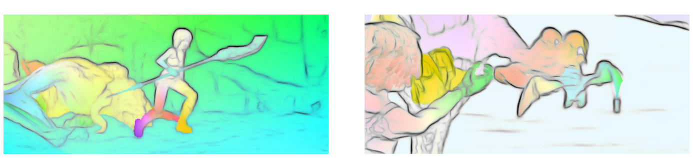

Figure 2: Image edges detected with SED [3] and ground-truth optical flow. Motion discontinuities appear most of the time at image edges.

The obtained interpolated field of correspondences is sufficiently accurate to be used as initialization of a one-level energy minimization. Our work thus suggests that there may be better initialization strategies than the well-established coarse-to-fine scheme. EpicFlow performs best on the challenging MPI-Sintel dataset and is competitive on Kitti and Middlebury.

Method	AEE	AEE-occ	s0-10	s10-40	s40+	Time
EpicFlow	**6.285**	**32.564**	**1.135**	**3.727**	**38.021**	**16.4s**
TF+OFM	6.727	33.929	1.512	3.765	39.761	~400s
DeepFlow	7.212	38.781	1.284	4.107	44.118	19s
S2D-Matching	7.872	40.093	1.172	4.695	48.782	~2000s
Classic+NLP	8.291	40.925	1.208	5.090	51.162	~800s
MDP-Flow2	8.445	43.430	1.420	5.449	50.507	709s

Table 1: Results on MPI-Sintel test set (final version). AEE-occ is the Average Endpoint Error on occluded areas. s0-10 is the AEE for pixels whose motions is between 0 and 10 px and similarly for s10-40 and s40+.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

Table 1 shows the results on the challenging MPI-Sintel dataset [2]. EpicFlow outperforms the state of the art with a gap of 0.5 pixel in AEE compared to the second best performing method, TF+OFM, and 1 pixel compared to the third one, DeepFlow [6]. In particular, we improve for both AEE on occluded areas and AEE over all pixels and for all displacement ranges. In addition, our approach is significantly faster than most of the methods, *e.g.* an order of magnitude faster than the second best. We can see in the examples of Figure 3 that EpicFlow better respects motion boundaries and is able to preserve small details. We provide in the paper an extensive evaluation of our method in terms of input matches and contours as well as a comparison to a standard coarse-to-fine scheme.

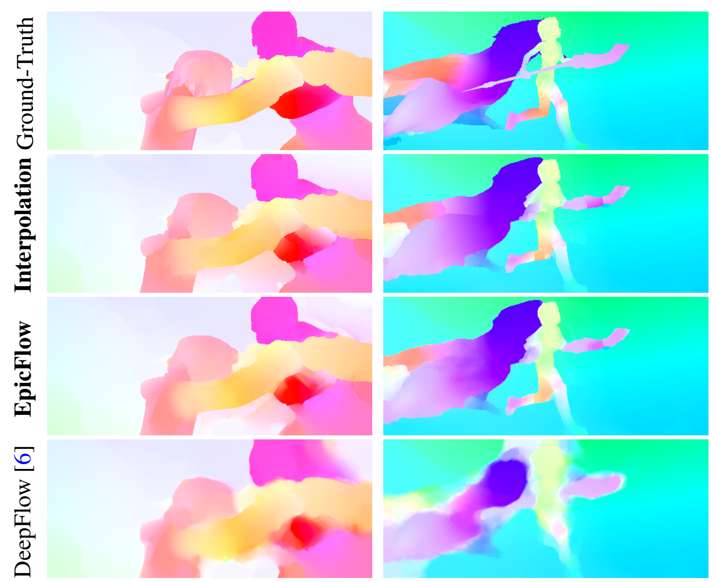

Figure 3: Example results.

[1] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *ECCV*, 2004.

[2] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In *ECCV*, 2012.

[3] P. Dollár and C. L. Zitnick. Structured forests for fast edge detection. In *ICCV*, 2013.

[4] M. Leordeanu, A. Zanfir, and C. Sminchisescu. Locally affine sparse-to-dense matching for motion and occlusion estimation. In *ICCV*, 2013.

[5] D. Sun, S. Roth, and M. J. Black. A quantitative analysis of current practices in optical flow estimation and the principles behind them. *IJCV*, 2014.

[6] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid. Deepflow: Large displacement optical flow with deep matching. In *ICCV*, 2013.