# Combinatorial Regularization of Descriptor Matching for Optical Flow Estimation

Benjamin Drayer
drayer@cs.uni-freiburg.de

Thomas Brox
brox@cs.uni-freiburg.de

Department of Computer Science,
Centre of Biological Signalling Studies (BIOSS),
University of Freiburg, Germany
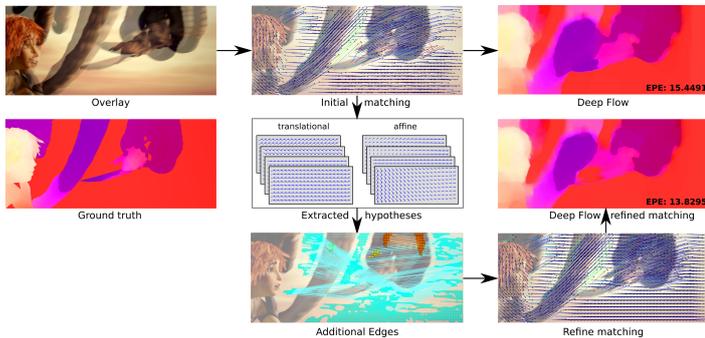
Figure 1: Outline of our method. The first row shows the overlaid input frames, the matching step (red arrows=high error, blue=low error) and the estimated optical flow using Deep Flow [5]. Starting from the initial matching, we extract translational and affine hypotheses. The 4-connected grid in the optimization/refinement step is enhanced by edges connecting similar regions. Our refinement improves at several position (note the fewer erroneous flow vectors and the denser field).

**Introduction** Dealing with large displacements has been the main research focus in the field of optical flow estimation in recent years [1, 5]. To estimate the large motion of small, detailed structures, a popular strategy is to pre-compute correspondences based on discriminative descriptors, such as HOG. While [1] relies on simple nearest neighbor matching to establish such correspondences, subsequent works have proposed more involved strategies to avoid wrong correspondences. The difference in performance between LDOF [1] and DeepFlow [5] is due to a more involved hierarchical matching strategy.

**Refinement** The contribution of the present paper is a refinement algorithm that further improves a given set of correspondences in several ways: (1) The refinement is computed densely (on a grid), which increases the number of matches. (2) Due to the explicit regularization, ambiguities in homogeneous regions or areas with repetitive texture are resolved. (3) Initially wrongly estimated outliers can be corrected. See Figure 1 for an outline of the algorithm.

We cast the refinement problem as a combinatorial energy minimization problem. The energy consists of three terms:

$$E(\mathcal{L}) = \underbrace{E_A(\mathcal{L}) + E_M(\mathcal{L})}_{\text{Data terms}} + \underbrace{E_S(\mathcal{L})}_{\text{Smoothness}} \quad (1)$$

The task is to find a labeling $\mathcal{L}$ that yields the minimal energy. The individual labels are affine motion hypotheses, derived with an alternating optimization scheme from the initial matching. For each hypothesis we assign a matching cost $E_M$ on how similar the displacement is to the initial guess (if there is one) and an appearance term $E_A$ that measures the similarity of the matched features. We regularize in the space of affine motion with the so-called flow continuity term [3]. While it is well-established to use a 4- or 8-connected neighborhood for discrete regularization, we propose to extend the set of edges by connecting homogeneous regions which likely belong together. In this way, the matching in enclosed homogeneous areas without sufficient characteristics is improved. Figure 2 gives an overview of the extended regularization.

**Results** The proposed refinement is a self-contained procedure and can be employed in any matching based optical flow algorithm. We show the improvement of our matches on the challenging Sintel dataset [2], where we evaluate the proposed refinement in the framework of LDOF, DeepFlow and EpicFlow (see Table 1). Qualitative results are shown in Figure 3.



Figure 2: Benefit of additional edges. When matching the first frame to the second (b), we compute regions with little texture, highlighted as black and colored areas in (c). Noise and small regions are discarded (marked as black). The remaining regions are connected if their appearance is similar. Connected regions have the same color. The bottom row shows the improvement from the initial LDOF [1] correspondences (d), over the refined matches (e) to and the refinement with the extended set of edges (f). The arrows are colored from blue to red (small to high error).
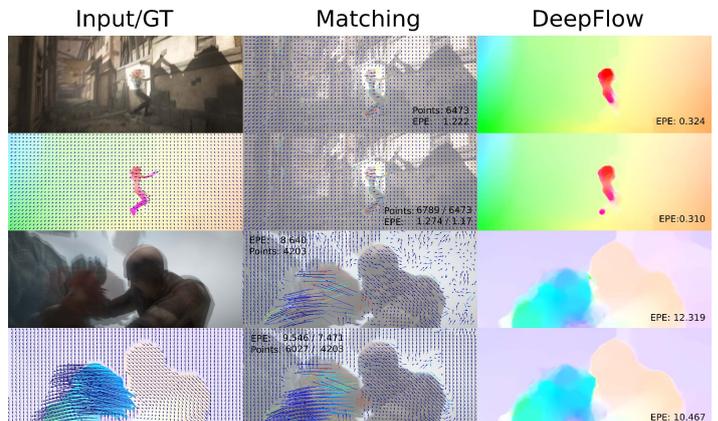


Figure 3: The upper line of each example shows the overlay of the input frames, the initial matching and the results of DeepFlow. The bottom line contains, groundtruth, our refined matching and the respective optical flow.

[1] T. Brox and J. Malik. Large displacement optical flow: descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513, 2011.

[2] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In *European Conf. on Computer Vision (ECCV)*, pages 611–625. Springer-Verlag, 2012.

[3] Y. Jiaolong and L. Hongdong. Dense, accurate optical flow estimation with piecewise parametric model. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[4] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid. EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[5] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid. DeepFlow: Large displacement optical flow with deep matching. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1385–1392. IEEE, 2013.

| LDOF [1] | LDOF+R | Deep [5] | Deep+R | Epic [4] | Epic+R |
|---|---|---|---|---|---|
| 6.026 | **5.616** | 4.022 | **3.852** | 3.566 | **3.497** |

Table 1: Comparison of the different optical flow methods with (+R) and without our regularization on the final pass of the Sintel training dataset. Throughout the methods, we get better results with the refined matches.