EVALUATION OF FLOWS AT THE OPTICAL LINE TERMINAL

Kayumova Shaxnoza Muxamajonovna

Elov J. B.

ANNOTATION

As a subfield of computer vision matures, datasets for quantitatively evaluating algorithms are essential to ensure continued progress. Many areas of computer vision, such as stereo, face recognition and object recognition, have challenging datasets to track the progress made by leading algorithms and to stimulate new ideas. Optical flow was actually one of the first areas to have such a benchmark, introduced by Barron et al. The field benefited greatly from this study, which led to rapid and measurable progress. To con tenue the rapid progress, new and more challenging datasets are needed to push the limits of current technology, revea where current algorithms fail, and evaluate the next gener ation of optical flow algorithms. Such an evaluation dataset for optical flow should ideally consist of complex real scenes with all the artifacts of real sensors (noise, motion blur, etc.). It should also contain substantial motion discontinuities and nonrigid motion. Of course, the image data should be paired with dense, subpixel-accurate, ground-truth flow fields.

The presence of nonrigid or independent motion makes collecting a ground-truth dataset for optical flow far harder than for stereo, say, where structured light or range scanning can be used to obtain ground truth. Our solution is to collect four different datasets, each satisfying a different subset of the desirable properties above. The combination of these datasets provides a basis for a thorough evaluation of current optical flow algorithms. Moreover, the relative performance of algorithms on the different datatypes may stimulate fur their research. In particular, we collected the following four types of data:

• Real Imagery of Nonrigidly Moving Scenes: Dense ground-truth flow is obtained using hidden fluorescent texture painted on the scene. We slowly move the scene, at each point capturing separate test images (in visible light) and ground-truth images with trackable texture (in UV light). Note that a related technique is being

used commercially for motion capture recently used certain wavelengths to hide ground truth in intrinsic images. Another form of hidden markers was also used in Ramnath et al to provide a sparse ground-truth alignment (or flow) of face images. Finally, Liu et al. recently proposed a method to obtain ground-truth using human annotation.

• Realistic Synthetic Imagery: We address the limitations of simple synthetic sequences such as Yosemite (Barron et al. by rendering more complex scenes with larger motion ranges, more realistic texture, independent motion, and with more complex occlusions.

• Imagery for Frame Interpolation: Intermediate frames are withheld and used as ground truth. In a wide class of application's such as video re-timing, novel-view generation, and motion-compensated compression, what is important is not how well the flow matches the ground-truth motion, but how well intermediate frames can be predicted using the flow.

• Real Stereo Imagery of Rigid Scenes: Dense ground truth is captured using structured light.

The data is then adapted to be more appropriate for optical flow by cropping to make the disparity range roughly symmetric.

We collected enough data to be able to split our collection into a training set (12 datasets) and a final evaluation set (12 datasets). The training set includes the ground truth and is meant to be used for debugging, parameter estimation, and possibly even learning. The ground truth for the final evaluation set is not publicly available (with the exception of the Yosemite sequence, which is included in the test set to allow some comparison with algorithms published prior to the release of our data).

We also extend the set of performance measures and the evaluation methodology of Barron et al. (1994) to focus attention on current algorithmic problems:

• Error Metrics: We report both average angular error and flow endpoint error (pixel distance). For image interpolation, we com pute the residual RMS error between the interpolated image and the ground-truth image. We also report a gradient normalized RMS error.

• Statistics: In addition to computing averages and standard deviations as in Barron et al. (1994), we also compute robustness measures and percentile-based accuracy measures.

• Region Masks: Following Scharstein and Szeliski, we compute the error measures and their statistics over certain masked regions of research interest. In particular, we compute the statistics near motion discontinuities and
intextureless regions.

Note that we require flow algorithms to estimate a dense flow field. An alternate approach might be to allow algorithms to provide a confidence map, or even to return a sparse or incomplete flow field. Scoring such outputs is problematic, however. Instead, we expect algorithms to gen tea flow estimate everywhere (for instance, using internal confidence measures to fill in areas with uncertain flow estimates due to lack of texture).

In October 2023 we published the performance of several well-known algorithms on a preliminary version of our data to establish the current state of the art. We also made the data freely available on the web at. Subsequently a large number of researchers have uploaded their results to our website and published papers using the data. A significant improvement in performance has already been achieved. In this paper we present both results obtained by classic algorithms, as well as results obtained since publication of our preliminary data. In addition to summarizing the overall conclusions of the currently uploaded results, we also examine how the results vary:

(1) across the metrics, statistics, and region masks,

(2) across the various datatypes and datasets,

(3) from flow estimation to interpolation, and

(4) depending on the components of the algorithms.

Optical flow estimation is an extensive field. A fully com prehensile survey is beyond the scope of this paper. In this related work section, our goals are: (1) to present a taxonomy of the main components in the majority of existing optical flow algorithms, and (2) to focus primarily on recent work and place the contributions of this work in the context of our taxonomy. Note that our taxonomy is similar to those of Stiller and Konrad for optical flow for stereo. For more extensive coverage of older work, the reader is referred to previous surveys such as those by Aggarwal and Nandhakumar, Barron et al. Otte and Nagel, Mitiche and Bouthemy, and Stiller and Konrad.

We first define what we mean by optical flow. Following taxonomy, the motion field is the 2D projection of the 3D motion of surfaces in the world, whereas the optical flow is the apparent motion of the brightness pat terns in the image. These two motions are not always the same and, in practice, the goal of 2D motion estimation is application dependent. In frame interpolation, it is prefer able to estimate apparent motion so that, for example, highlights move in a realistic way. On the other hand, in applications where the motion is used to interpret or reconstruct the 3D world, the motion field is what is desired.

In this paper, we consider both motion field estimation and apparent motion estimation, referring to them collectively as optical flow. The ground truth for most of our datasets is the true motion field, and hence this is how we define and evaluate optical flow accuracy. For our interpolation datasets, the ground truth consists of images captured at an intermediate time instant. For this data, our definition of optical flow is really the apparent motion.

We do, however, restrict attention to optical flow algorithms that estimate a separate 2D motion vector for each pixel in one frame of a sequence or video containing two or more frames. We exclude transparency which requires multiple motions per pixel. We also exclude more global representations of the motion such as parametric motion estimates

REFERENCES

1. Rakhmanov, A.T.; Ruziboev, O.B. To the solution of the classification problem for three classes. Science and world 2015, 2(18), 21–25.

2. Nishanov, A.X.; Ruzibaev, O.B.; Tran, N.H. Modification of decision rules "ball Apolonia" the problem of classification. Int. Conf. on Information Science and Communications Technologies ICISCT 2016, 1–3.
3. Mahajan, P.: Applications of pattern recognition algorithm in health and medicine. Int. J. Eng. Comput. Sci. 5, 16580–16583 (2016)
4. Ashok, B., Aruna, P.: Comparison of feature selection methods for diagnosis of cervical cancer using SVM classifier. J. Eng. Res. Appl. 6(1), 94–99 (2016)
5. Nishanov, A.K., Akbaraliev, B.B., Samandarov, B.S., Akhmedov, O.K., Tajibaev, S.K.: An algorithm for classification, localization and selection of informative features in the space of politypic data. Webology 17(1), 341–364 (2020)