Course 049064: Variational Methods in Image Processing

Notes 8: Optical Flow

Guy Gilboa

1 Basic Model

1.1 Background

Optical flow is a fundamental problem in computer vision. The general goal is to find a displacement field between two similar images. This is related to different vision tasks, three main ones are:

1. *Motion Estimation.* This is the origins of the optical flow problem. When processing and analyzing videos, crucial information is which objects moved and where to. This allows tracking of people, cars or moving targets, analyzing gestures, estimating camera motion etc.

2. *Image Registration.* Transforming a set of images into a sinlge coordinate system. Very useful in medical imaging, automatic target recognition and to process satellite images. The registration allows easy integration and comparison of the data. Can be single or multi modal (various sensors or image acquisition technologies).

3. *Stereo and 3D Reconstruction.* Computing the geometry of shapes in the image based on several images, either from multiple cameras or from a sequence taken from a moving camera.

Here we will focus on the optical flow problem for motion estimation.

1.2 The optical flow equation

We assume to have a video sequence $f(x, y, t)$, where (x, y) are the spatial coordinates and t is the time. Note that here t stands for actual time of the sequence (and is not an artificial evolution parameter). For simplicity we consider two images, sampled at a time interval dt: $f(x, y, t)$ and $f(x, y, t + dt)$. The optical flow assumption is of *gray level constancy*. That is - the objects can change location, but their color stays the same. Having a displacement vector $(\Delta x, \Delta y)$ for each pixel we can write:

$$f(x, y, t) = f(x + \Delta x, y + \Delta y, t + dt). \tag{1}$$

Taking a first order Taylor approximation we have:

$$f(x+\Delta x, y+\Delta y, t+dt) = f(x, y, t)+\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y+\frac{\partial f}{\partial t}dt+\text{Higher-order-terms.} \tag{2}$$

Writing the displacement vector as $v = (\frac{\Delta x}{dt}, \frac{\Delta y}{dt}) = (v_1(x, y), v_2(x, y))$, we get the first order approximation of the flow, known as the optical-flow-equation (or optical-flow-constraint):

$$f_x v_1 + f_y v_2 + f_t = 0. \tag{3}$$

Note that this optical flow model is very simple and is fulfilled only partly in real-world scenarios. There are many reasons why in practice this equation will not model well some motions in the image. The mains deviations from the intensity constancy assumption are due to:

- *Occlusions.* Objects which appear in the first image can be occluded in the second, and vice versa. Therefore pixel intensities can "disappear" or "pop up" between frames.
- *Lighting.* The lighting changes in the scene spatially, therefore the pixel intensity changes when the object moves. Also lighting can change in time (turning on/off lights, clouds moving etc.).
- *Shadows.* This is a specific part of light change, which is very common and is hard to avoid. For instance, the shadow of a person casted on the floor is moving, although the floor, naturally, is still.
- *Normal angle change.* The angle of the normal of a moving object changes with respect to the camera. Thus the intensity can change.
- *Scaling.* The resolution and camera parameters are not taken into account. For example, when a textured object approaches the camera - a single pixel can "become" two pixels in the next frame with different intensities.

- *Movement at boundaries.* For still camera - moving objects come into the frame and out of the frame. For moving camera, or zooming in or out, in addition, the entire scene can change at the boundaries (also of still objects).

Under-determined problem: Another problem of solving the optical flow problem is that the optical-flow-equation (3) gives a single equation for every pixel, whereas the vector flow v has two components for every pixel. Therefore the problem is under-determined and we need to resort to spatial regularity constraints.

The aperture problem: For constant regions in the image we get a large ambiguity, where one cannot determine local motion. The equations become degenerate, with many possible solutions. In these cases one needs to resort to additional assumptions of significant spatial regularity.

1.3 Early attempts

There has been two different distinct directions to solve the problem, a local approach suggested by Lucas and Kanade in [7] and a global approach suggested by Horn and Schunk in [6].

Lucas-Kanade: They assumed the flow is essentially constant in a local neighborhood of the pixel. For instance, for two pixels, one can write Eq. (3) for each pixel and assume (v_1, v_2) are the same. Therefore the problem can be theoretically solved. To increase robustness they increased the window size and solved a least-square problem. In the case of large flat regions - there was still large ambiguity.

Horn-Schunck: They suggested a global variational approach. To solve the optical flow problem the following functional was minimized:

$$E_{HS} = \int_\Omega (f_x v_1 + f_y v_2 + f_t)^2 + \alpha(|\nabla v_1|^2 + |\nabla v_2|^2) dx. \tag{4}$$

This is a convex problem which can be solved numerically, for example by a gradient descent based on the Euler-Lagrange equations with respect to v_1 and v_2:

$$\begin{aligned} f_x(f_x v_1 + f_y v_2 + f_t) - \alpha \Delta v_1 &= 0, \\ f_y(f_x v_1 + f_y v_2 + f_t) - \alpha \Delta v_2 &= 0. \end{aligned} \tag{5}$$

As noted in [4], this type of solution has a filling-in effect: at location with $|\nabla f| \approx 0$ no reliable local flow estimate is possible (aperture problem) but the regularizer $|\nabla v_1|^2 + |\nabla v_2|^2$ fills in information from the neighborhood, resulting in a dense flow. This is a clear advantage over local methods.

In the next section we will follow this approach.

2 Modern optical flow techniques

There has been many attempts to solve the various problems encountered in optical flow, using various regularizers and different models, for some of them see [1, 4, 9, 2, 8, 5].

We will present in more depth one of the more successful solutions, which is simple enough to understand, yet performs very well and was considered state-of-the-art for a long duration since its publication in 2004.

2.1 Brox et al model

This model was suggested by Brox-Bruhn-Papenberg-Weickert (BBPW) in [3]. The total energy is composed of the standard general two terms:

$$E_{BBPW}(v) = E_{Data} + \alpha E_{Smooth}. \tag{6}$$

Let us define $\mathbf{w} = (v_1, v_2, 1)$, $\mathbf{x} = (x, y, t)$ and for simplicity we assume the difference in time between frames is a unit time $dt = 1$. We can also define a spatio-temporal gradient by

$$\nabla_3 = (\partial_x, \partial_y, \partial_t).$$

Let

$$\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}.$$

The model is essentially an L^1 model of the data and smoothness terms.

The data term consists of gray-level constancy as well as gradient constancy assumption:

$$E_{Data}(v) = \int_\Omega \Psi(|f(\mathbf{x}+\mathbf{w}) - f((\mathbf{x}))|^2 + \gamma|\nabla f(\mathbf{x}+\mathbf{w}) - \nabla f((\mathbf{x}))|^2)d\mathbf{x}. \tag{7}$$

We assume piece-wise smoothness of the flow field. Therefore a total-variation approximation of the flow field is minimized:

$$E_{Smooth}(v) = \int_\Omega \Psi(|\nabla_3 v_1|^2 + |\nabla_3 v_2|^2)d\mathbf{x}. \tag{8}$$

For the quite sophisticated numerical methods for solving this model - see the paper [3].

References

[1] G Aubert, R Deriche, and P Kornprobst. Computing optical flow via variational techniques. *SIAM Journal on Applied Mathematics*, 60(1):156–182, 1999.

[2] Simon Baker, Daniel Scharstein, JP Lewis, Stefan Roth, Michael J Black, and Richard Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, 2011.

[3] Thomas Brox, Andrés Bruhn, Nils Papenberg, and Joachim Weickert. High accuracy optical flow estimation based on a theory for warping. In *Computer Vision-ECCV 2004*, pages 25–36. Springer, 2004.

[4] Andrés Bruhn, Joachim Weickert, and Christoph Schnörr. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005.

[5] Daniel Cremers and Stefano Soatto. Motion competition: A variational approach to piecewise parametric motion segmentation. *International Journal of Computer Vision*, 62(3):249–265, 2005.

[6] Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1):185–203, 1981.

[7] Bruce D Lucas, Takeo Kanade, et al. An iterative image registration technique with an application to stereo vision. In *IJCAI*, volume 81, pages 674–679, 1981.

[8] Deqing Sun, Stefan Roth, and Michael J Black. Secrets of optical flow estimation and their principles. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2432–2439. IEEE, 2010.

[9] Christopher Zach, Thomas Pock, and Horst Bischof. A duality based approach for realtime tv-l 1 optical flow. In *Pattern Recognition*, pages 214–223. Springer, 2007.