Optical Flow

Prof. H. Ișıl Bozma

Electrical and Electronics Engineering

2010
Contents

12 Optical Flow
12.1 Time-varying Images ................................................. 1
12.2 Motion Field .......................................................... 2
12.3 Optical Flow .......................................................... 3
12.4 Optical Flow Equation ............................................... 3
12.5 Correspondence Problem ............................................. 4
12.6 Differential Methods .................................................. 5
List of Figures

12.1 Object motion. ......................................................... 2
12.2 Motion field of a rigid object. ................................. 2
12.3 Focus of expansion. ............................................. 3
12.4 Barber’s pole. .................................................... 3
12.5 Aperture problem. ............................................... 4
Chapter 12

Optical Flow

12.1 Time-varying Images

So far, images have had only two dimensions – namely the image plane coordinates. However, images also change in time which implies a third dimension – the time axis. Here, instead of a single image, we have a sequence of image. Due to the imaging devices, images are sampled over time which implies a discrete nature.

Time-varying images present both new possibilities and difficulties. For example, camouflaged objects are only easily seen when they move. Moreover, the relative sizes and position of objects are more easily determined when the objects move. Even simple image differencing provides an edge detector for the silhouettes of texture-free objects moving over any static background.

There are different types of motion to consider:

- Motion of objects only (static camera)
- Motion of camera (ego motion) – no motion of the objects
- Both objects and camera moving

Since motion is relative, one would expect the results of these different types of motion to be same. However, this is not the case. For example, if the objects move relative to the illumination source, specularity or shadow effects may be changing. For now, such complications will be ignored.

The analysis of visual motion is divided into two stages:

- The measurement of the motion,
- Discontinuities can help segmenting the scene into distinct objects and to extract three dimensional information about the shape and motion of the objects.
12.2 Motion Field

The image of an object moving in front of a camera causes a corresponding change in the image. Consider an object point \( p_o \) moving with a velocity \( v_o \). The corresponding image point \( p_i \) moves on the image plane with velocity vector \( v_i \) as seen in Fig. 12.1. The set of these velocity vectors is referred to as the motion field. In case of rigid body translations and rotations, the motion field is continuous except at the silhouette boundaries of objects as seen in Fig. 12.2.

Figure 12.2 Motion field of a rigid object.

In the case of pure camera translation:

- **Focus of expansion/contraction:** (FOE/FOC) The image point from which (or towards which) all motion vectors radiate – namely the point of divergence (or convergence) of all motion field vectors. (Fig. 12.3)

- The direction of motion is along the projection ray through FOE.

- In the case of divergence \( \implies \) Forward motion of the camera.

- In the case of convergence \( \implies \) Backwards motion.

- If the axis of camera translation is taken as the camera baseline in stereo, then every projection of a fixed scene point must translate along an epipolar line, and all such lines converge at the epipole, which is just the FOE.
12.3 Optical Flow

- *Optical flow* is the apparent motion of brightness patterns observed when a camera is moving relative to the objects being observed.

- Generally, optical flow corresponds to the motion field, but not always as seen in Fig. 12.4. However, in general, such cases are unusual, and we will assume identicalness.

![Figure 12.4 Barber’s pole.](image)

- *Aperture Problem:* Can measure the component of optical flow that is in the direction of the intensity gradient! (Fig. 12.5)

12.4 Optical Flow Equation

Consider the intensity function $I : X \times R \rightarrow C$ as $I(x_1, x_2, t) \in C$. The rate of change of $I$ wrt to time can be expressed as:

$$\frac{dI}{dt} = \frac{\partial I}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial I}{\partial x_2} \frac{dx_2}{dt} + \frac{\partial I}{\partial t}$$  \hspace{1cm} (12.0)

If we assume time-invariant illumination and no shadow effects, then

$$\frac{dI}{dt} = 0$$  \hspace{1cm} (12.0)
12.5. CORRESPONDENCE PROBLEM

which implies
\[ D_{x_1} I \dot{x}_1 + D_{x_2} I \dot{x}_2 + D_t I = 0 \] (12.0)
which can be rewritten as
\[ \begin{bmatrix} D_{x_1} I \\ D_{x_2} I \end{bmatrix}^T \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = -D_t I \] (12.0)

As seen, there is information only regarding the component of the optical flow in the direction of intensity gradient. If possible, we should choose image points that has significant intensity gradient in more than one direction – for example corner points.

12.5 Correspondence Problem

The ultimate aim of analyzing motion in images is reconstruct the 3D motion and structure of the observed world. To do so, we must first tackle the correspondence problem: Which pixels of an image correspond to which pixels of the next frame of the image sequence?

There are two main approaches:

1. **Differential methods**: Compute the optical flow and use its geometrical properties to deduce three dimensional information about the scene and the motion.
   - Optical flow needs to be computed at regular pixel grid while enforcing some smoothness constraint of the optical flows over neighbouring pixels,
   - Leads to dense 3D structure

2. **Matching methods**: Convert the motion problem to a stereo problem and find the correspondence between a number of (sparse) feature points (e.g. corners) in the image at time \( t \) to the image at time \( t + \delta t \).
   - Leads to sparse 3D structure
12.6 Differential Methods

- Formulate as solving the optical flow equation for surface orientation variables $u$ and $v$ where $u = \dot{x}_1$ and $v = \dot{x}_2$:
  \[
  \int \int D_{x_1}Iu + D_{x_2}Iv + D_tI \, dx_1 dx_2
  \]

- Underconstrained $\implies$ No unique solution

- Ensure regularity while requiring agreement of Image Irradiance Equation over the region of interest

  \[
  \min \int \int u_{x_1}^2 + u_{x_2}^2 + v_{x_1}^2 + v_{x_2}^2 + \lambda(D_{x_1}Iu + D_{x_2}Iv + D_tI)^2 \, dx_1 dx_2
  \]

Letting

\[
F(u, v, \lambda) = u_{x_1}^2 + u_{x_2}^2 + v_{x_1}^2 + v_{x_2}^2 + \lambda(D_{x_1}Iu + D_{x_2}Iv + D_tI)^2
\]

The corresponding Euler equations are:

\[
D_u F - \frac{\partial}{\partial x_1} D_{ux_1} - \frac{\partial}{\partial x_2} D_{ux_2} = 0
\]

\[
D_v F - \frac{\partial}{\partial x_1} D_{vx_1} - \frac{\partial}{\partial x_2} D_{vx_2} = 0
\]

\[
\nabla^2 u = -\lambda(D_{x_1}Iu + D_{x_2}Iv + D_tI)D_{x_1}I
\]

\[
\nabla^2 v = -\lambda(D_{x_1}Iu + D_{x_2}Iv + D_tI)D_{x_2}I
\]

where

\[
\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}
\]  \hspace{1cm} (12.-4)

The discretized version is

\[
E(u, v) = \sum_i \sum_j s_{ij} + \lambda r_{ij}
\]

\[
s_{ij} = \frac{1}{4}((u_{i+1,j} - u_{i,j})^2 + (u_{i,j+1} - u_{i,j})^2 + (v_{i+1,j} - v_{i,j})^2 + (v_{i,j+1} - v_{i,j})^2)
\]

\[
r_{ij} = (D_{x_1}Iu + D_{x_2}Iv + D_tI)^2
\]

Differentiating wrt to $u_{ij}$ and $v_{ij}$ and setting the resulting equations to 0,

\[
u_{ij}^{n+1} = \frac{D_{x_1}Iw_{ij}^n + D_{x_2}Iv_{ij}^n + D_tI}{1 + \lambda(D_{x_1}I^2 + D_{x_2}I^2)} D_{x_1}I
\]

\[
v_{ij}^{n+1} = \frac{D_{x_1}Iw_{ij}^n + D_{x_2}Iv_{ij}^n + D_tI}{1 + \lambda(D_{x_1}I^2 + D_{x_2}I^2)} D_{x_2}I
\]
The computation for the average can be computed using the stencil:

\[
\begin{array}{ccc}
1 & 4 & 1 \\
4 & 0 & 4 \\
1 & 4 & 1 \\
\end{array}
\]

(12.9)

In summary, the new value of \((u, v)\) at a point is equal to the average of the surrounding values minus an adjustment in the direction of the brightness gradient – using estimates of the values \(D_{x1}I, D_{x2}I, and D_{t}I\). These can be computed by taking local averages in neighbourhoods about the grid point \((i; j; k)\).

Which one of these approaches is more suitable?

- Depends on the type of images.
- If many prominent image corners are present in your images and if these corners can be detected and matched \(\Rightarrow\) Matching Method
- Otherwise \(\Rightarrow\) Differential Method
Bibliography