PROJECT #3 Optical Flow
By Jie Zou
November 17, 2001

1. Introduction

Optical flow is a vector field in the image that represents an approximation of the image motion field. It represents motion field in the direction of the image intensity gradient. In this project, we have two sets of 5 consecutive frames. One is rotation of a synthetic sphere. The other is movement of two humans. I have computed the optical flow for the central frame of each sequence, and determined the moving direction of the objects.

2. Mathematic discussion of optical flow and its estimation method

The constancy of the apparent brightness of the observed scene can be written as the stationarity of the image brightness I over the time, so: $\frac{dI}{dt} = 0$. This constraint is completely satisfied under 1) the motion is translation; 2) illumination direction is parallel to the angular velocity for lambertian surface.

I is a function of (x, y), which, in turn, are function of time t. Hence: $\frac{dI}{dt} = \frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$, where $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ represent spatial intensity gradient while $\frac{\partial I}{\partial t}$ represents temporal intensity gradient. This is the image brightness constancy equation $(\nabla I)^t v + I_t = 0$. ∇I is called image intensity gradient. This equation can be used to estimate the motion field v. This equation has no constraint on v when it is orthogonal to ∇I, so this equation can only determine motion flow component in the direction of the intensity gradient, i.e., the projection of motion field v in the gradient direction. This special motion field is called *optical flow*. Optical flow is always parallel to image gradient.

For each image point p and a $N \times N$ neighborhood R, where p is the center, assume every point in the neighborhood has the same optical flow v, we have $\nabla^t I(x,y)v(x,y) + I_t(x,y) = 0 \quad (x,y) \in R$. $v(x,y)$ can be estimated via $e^2 = \sum_{(x,y)\in R} \left(\nabla^t I(x,y)v(x,y) + I_t(x,y)\right)^2$. The least-squares solution to $v(x,y)$ is

$$v(x,y) = \left(A^t A\right)^{-1} A^t b, \text{ where } A = \begin{bmatrix} \nabla^t I(x_1, y_1) \\ \nabla^t I(x_2, y_2) \\ \cdots \\ \nabla^t I(x_N, y_N) \end{bmatrix}, \text{ and } b = -\begin{bmatrix} I_t(x_1, y_1) \\ I_t(x_2, y_2) \\ \cdots \\ I_t(x_N, y_N) \end{bmatrix}.$$

Besides assuming brightness constancy while objects are in motion, we can assume smoothness constraint on the motion field, i.e., motion field projections in x, y, and t remain the same for a small neighborhood. Mathematically, these constraints can be formulated as follows: $\frac{d^2 I}{dtdx} = 0$, $\frac{d^2 I}{dtdy} = 0$, $\frac{d^2 I}{dtdt} = 0$. Then we have additional optical constraints:

$$v_x I_{xx} + v_y I_{yx} + I_{tx} = 0$$
$$v_x I_{xy} + v_y I_{yy} + I_{ty} = 0$$
$$v_x I_{xt} + v_y I_{yt} + I_{tt} = 0 .$$

Totally, we have four optical flow constraints:

$$v_x I_x + v_y I_y + I_t = 0$$
$$v_x I_{xx} + v_y I_{yx} + I_{tx} = 0$$
$$v_x I_{xy} + v_y I_{yy} + I_{ty} = 0$$
$$v_x I_{xt} + v_y I_{yt} + I_{tt} = 0 .$$

We can therefore solve $v = (v_x, v_y)$ using a linear least squares method:

$$v = (A^t A)^{-1} A^t b , \qquad (1)$$

where $A = \begin{pmatrix} I_x & I_y \\ I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \\ I_{tx} & I_{ty} \end{pmatrix}$, and $b = -\begin{pmatrix} I_t \\ I_{xt} \\ I_{yt} \\ I_{tt} \end{pmatrix}$.

Traditional approach to compute intensity derivatives involves numerical approximation of continuous differentiations. We compute image derivatives analytically using a cubic facet model to obtain an analytical and continuous image intensity function that approximates image surface at time (x, y, t).

Assume the gray level pattern of each small block in an image sequence is ideally a canonical 3D cubic polynomial of x, y, t:

$$\begin{aligned} I(x, y, t) = {} & a_1 + a_2 x + a_3 y + a_4 t + a_5 x^2 + a_6 xy + a_7 y^2 + a_8 yt + a_9 t^2 + a_{10} xt \\ & + a_{11} x^3 + a_{12} x^2 y + a_{13} xy^2 + a_{14} y^3 + a_{15} y^2 t + a_{16} yt^2 + a_{17} t^3 + a_{18} x^2 t + a_{19} xt^2 + a_{20} xyt \end{aligned} \qquad (2)$$

The solution for coefficients $a = (a_1, \quad a_2, \quad \cdots \quad a_{20})^t$ can be solved using least square: $a = (D^t D)^{-1} D^t I$, (3) where $D = \begin{pmatrix} 1 & x_1 & y_1 & t_1 & \cdots & x_1 y_1 t_1 \\ 1 & x_1 & y_1 & t_2 & \cdots & x_1 y_1 t_2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_1 & y_2 & t_2 & \cdots & x_1 y_2 t_1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_X & y_Y & t_T & \cdots & x_X y_Y t_T \end{pmatrix}$, and $I = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{pmatrix}$.

Image derivatives are readily available from the cubic facet model. Substituting a into above A and b, we get: $A = \begin{pmatrix} a_2 & a_3 \\ 2a_5 & a_6 \\ a_6 & 2a_7 \\ a_{10} & a_8 \end{pmatrix}$, $b = -\begin{pmatrix} a_4 \\ a_{10} \\ a_8 \\ 2a_9 \end{pmatrix}$. Then, we can use equation (1) to compute optical flow.

3. Experiment procedure and results

Based on the above mathematic analysis, we use the following algorithm to estimate optical flow:

1. Select an image as central frame (the 3rd frame of 5 frames)
2. For each pixel (excluding the boundary pixels) in the central frame:
   a. Perform a cubic facet model fit using equation (2) and obtain the 20 coefficients using equation (3).
   b. Derive image derivatives using the coefficients and the A matrix and b vector.
   c. Compute image flow using equation (1).
   d. Mark each point with an arrow indicate its flow if its flow magnitude is larger than a threshold.

I have applied the above algorithm on both image sequences, and the optical flow for both sequences are attached at the end of this report.

For the sphere rotation, it is a synthetic sequence, so the optical flow clearly shows that the sphere is rotating along an axis.

For the human moving sequence, since it is a real moving scene, the optical flow is not as clear as the above synthetic sequence. But, from the optical flow, we still can observe that there are two moving humans. In the left, there is a human moving to the left, and in the right, there is a human moving away from the camera. Based on the length of the arrow (the magnitude of optical flow), we can say that the left human is moving faster than the right human.

During the implementation of the optical flow algorithm presented above, I noticed that for the real moving scene (the human moving sequence), the magnitude of some computed optical flow is very large compare to the other "correct" optical flow. I think these high-magnitude optical flows are the "noise". There are two reasons: one is it is possible that $A^t * A$ is also nearly singular, and this may lead to computation instability; the other is due to the image noise of original frames. In order to get rid of these high-magnitude optical flows (since they are clearly not the true optical flow), I set one threshold. If the magnitude of optical flow is higher than the threshold, I set the optical flow to 0. For the moving human sequence, I set the following threshold: $v_x + v_y \leq 6$. For the rotation sphere, since it is a synthetic sequence, I didn't set any threshold.

4. Summary

In this project, I have implemented an algorithm to compute the optical flow for two sequences of frames. One is a synthetic sphere rotation, and the other is human moving scene. The computed optical flow can indicate how the objects are moving. One thing we need to take care is that for the real moving scene, due to the noise of the original images, the computed optical flow may not be "correct", and we should filter out some obvious incorrect optical flow in order to let the image of optical flow clearly show how the object is moving.