Computer Vision Project 3
Optical Flow Estimation

Zhiwei, Zhu

1 Introduction

If we take a series of images in time, and there are moving objects in the scene, or perhaps the camera is itself on some moving vehicle, useful information about what the image contains can be obtained by analyzing and understanding the difference between images caused by the motion.

The goal of motion analysis is to characterize the relative motion and use it as a visual cue for object detection, scene segmentation or 3D structure reconstruction. Motion is important since it represents spatial changes over time, which is crucial for understanding the dynamic world.

Studying the motion in detail, we can answer such questions as

1. How many moving objects are there?
2. Which directions are they moving in?
3. Whether they are undergoing linear or rotational motion?
4. How fast they are moving?
5. What are the structures of the moving objects?

In this project, we are going to calculate a new function called Optical Flow for an image sequence, which represents an approximation of the image motion field.

2 Optical Flow

Optical flow is a vector field in the image to characterize the motion field for each pixel. Specifically, for every pixel, a velocity vector (v_x, v_y) is found to represent

1. how quickly that pixel is moving across the image,
2. the direction it is moving.

Let us suppose that the intensity of pixel (x, y) at time t is given by $I(x, y, t)$, where the intensity is now a function of time, t, as well as of x and y.

At a point a small distance away, and a small time later, the intensity is

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + \ldots$$

where the dots stand for higher order terms.

Now, suppose that part of an object is at a position (x, y) in the image at a time t, and that by a time dt later it has moved through a distance (dx, dy) in the image.

Furthermore, let us suppose that the intensity of that part of the object is just the same in our image before and afterwards.

Provided that we are justified in making this assumption, we then have that

$$I(x+dx, y+dy, t+dt) = I(x, y, t)$$

and so

$$\frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + \ldots = 0$$

However, dividing through by dt, we have that

$$\frac{dx}{dt} = u$$
$$\frac{dy}{dt} = v$$

as these are the speeds the object is moving in the x and y directions respectively. Thus, in the limit that dt tends to zero, we have

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v = -\frac{\partial I}{\partial t} \tag{1}$$

which is called the optical flow constraint equation.

Now, $\frac{\partial I}{\partial t}$ at a given pixel is just how fast the intensity is changing with time, while $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ are the spatial rates of change of intensity, i.e. how rapidly intensity changes on going across the picture, so all three of these quantities can be estimated for each pixel by considering the images.

Finally, this equation can be used to estimate the optical flow $V = (u, v)$. Note this equation has no constraint on V when it is orthogonal to ∇I, which is the image intensity gradients $(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$. Or, even if the image gradients are zeros, which means $\nabla I = (0, 0)$, then we can not estimate the optical flow.

3 Optical Flow Estimation

To estimate the optical flow, we need an additional constraint since equation 1 only provides one equation for 2 unknowns.

For each image point p and a $N \times N$ neighborhood R, where p is the center, we assume that every point in the neighborhood has the same optical flow V. In here, we should note that this is a smoothness constraint and may not hold near the edges of the moving objects. From the equation

$$\nabla^t I(x,y)V(x,y) + I_t(x,y) = 0, (x,y) \in R$$

$V(x,y)$ can be estimated via

$$\epsilon^2 = \sum_{(x,y)\in R} \left(\nabla^t I(x,y)V(x,y) + I_t(x,y)\right)^2$$

Then, the least squares solution to $V(x,y)$ is

$$V(x,y) = (A^t A)^{-1} A^t b$$

where

$$A = \begin{pmatrix} \nabla^t I(x_1, y_1) \\ \nabla^t I(x_2, y_2) \\ \vdots \\ \nabla^t I(x_N, y_N) \end{pmatrix}$$

and

$$b = -\begin{pmatrix} I_t(x_1, y_1) \\ I_t(x_2, y_2) \\ \vdots \\ I_t(x_N, y_N) \end{pmatrix}$$

4 Additional Optical Flow Constraints

Besides assuming brightness constancy while objects are in motion, we can assume smoothness constraints 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:

$$\frac{d^2 I}{dt dx} = 0,$$
$$\frac{d^2 I}{dt dy} = 0,$$
$$\frac{d^2 I}{dt dt} = 0$$

Applying them to equation 1 yields three additional optical constraints:

$$V_x I_{xx} + V_y I_{yx} + I_{tx} = 0$$
$$V_x I_{xy} + V_y I_{yy} + I_{ty} = 0$$
$$V_x I_{xt} + V_y I_{yt} + I_{tt} = 0$$

The four optical flow constraints are

$$V_x I_x + V_y I_y + I_t = 0$$
$$V_x I_{xx} + V_y I_{yx} + I_{tx} = 0$$
$$V_x I_{xy} + V_y I_{yy} + I_{ty} = 0$$
$$V_x I_{xt} + V_y I_{yt} + I_{tt} = 0$$

So, this yields four equations for two unknowns $V = (V_x, V_y)$.
Therefore, they can be solved using a linear least squares method by minimizing

$$\|AV - b\|^2$$

where

$$A = \begin{pmatrix} I_x & I_y \\ I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \\ I_{tx} & I_{ty} \end{pmatrix}$$

and

$$b = -\begin{pmatrix} I_t \\ I_{xt} \\ I_{yt} \\ I_{tt} \end{pmatrix}$$

Then the optical flow V can be estimated by

$$V = A^t A^{-1} A^t b$$

5 Computing Image Derivatives

In order to compute the optical flow, we have to obtain the intensity derivatives first. Traditional approach to compute intensity derivatives involves numerical approximation of continuous differentiations. We propose to compute image derivatives analytically using a cubic facet model to obtain an analytical and continuous image intensity function that approximates image surface at (x, y, t). This yields more robust and accurate image derivatives estimation due to noise suppression via smoothing by function approximation.

5.1 Cubic Facet Model

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

$$I(x,y,t) = a_1 + a_2x + a_3y + a_4t + a_5x^2 + a_6xy + a_7y^2 + a_8yt + a_9t^2 + a_{10}xt + a_{11}x^3 + a_{12}x^2y + a_{13}xy^2 + a_{14}y^3 + a_{15}y^2t + a_{16}yt^2 + a_{17}t^3 + a_{18}x^2t + a_{19}xt^2 + a_{20}xyt$$

The solution for coefficients $a = (a_1, a_2, ..., a_{20})^t$ in the Least-squares sense minimizes $\|Da - J\|^2$ and is expressed by

$$a = (D^{'}D)^{-1}D^{'}J$$

where

$$D = \begin{pmatrix} 1 & x_1 & y_1 & t_1 & \dots & x_1y_1t_1 \\ 1 & x_1 & y_1 & t_2 & \dots & x_1y_1t_2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_1 & y_2 & t_2 & \dots & x_1y_2t_1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_X & y_Y & t_T & \dots & x_Xy_Yt_T \end{pmatrix}$$

and

$$J = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{pmatrix}$$

In which I_n is the intensity value at (x_i, y_i, t_k), X, Y are the column and row numbers of the pixel block, and K is the image frame number.

While performing surface fitting, the surface should be centered at the pixel (voxel) being considered and use a local coordinate system, with the center as its origin. Therefore, for a $3 \times 3 \times 3$ neighborhood, the coordinates for $x, yandt$ are: $-1\ 0\ 1$, $-1\ 0\ 1$ and $-1\ 0\ 1$ respectively.

Image derivatives are readily available from the cubic facet model. Substituting a_i into yields the OFCE's we actually use:

$$A = \begin{pmatrix} a_2 & a_3 \\ 2a_5 & a_6 \\ a_6 & 2a_7 \\ a_{10} & a_8 \end{pmatrix}, b = \begin{pmatrix} a_4 \\ a_{10} \\ a_8 \\ 2a_9 \end{pmatrix}$$

6 Optical Flow Estimation Algorithm

First, we are given a sequence of N frames, typically N is equal to 5. Let Q be a square region of $L \times L$, typically, L is equal to 5.

Finally, we summarize the steps for optical flow estimation using facet method as follows:

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

   Set the optical flow vectors to zero for locations where matrix A^tA is singular.

7 Experiments

Synthetic and real image sequences are used to test our optical flow estimation algorithm. In the following, we will show the experiment results in detail.

1. Synthetic Data

   The first synthetic image sequence contains a ball which is rotating counter-clock-wisely. Figure 1 shows the estimated optical flows of the central image in this image sequence.

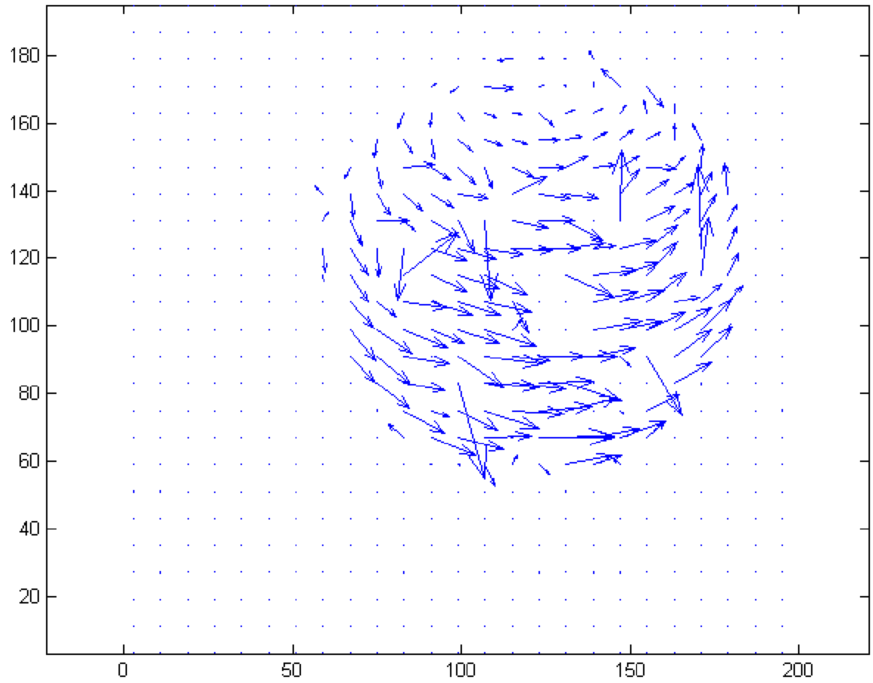


Figure 1. The estimated optical flow indicated by the arrows.

From Figure 1, we can see that for most of the pixels in the ball, the estimated optical flows are almost the same as the motion of the ball we observed. Also, we can see that for some pixels in the ball, the estimated optical flows are incorrect, but they can be neglected. Further we can see that the pixels which have incorrectly estimated optical flows are located in the dark region of the ball. Usually, there dark region has zero image gradient. According to the equation 1, because the image gradients are zeros, which means

$\nabla I = (0,0)$, then we can not estimate the optical flow. So, for these pixels, the estimated optical flows are not accurate.

The second image sequence contains a rectangle object moving from left to right. The estimated optical flows are shown in the Figure 2. We can that for most of estimated optical flows are same as the observed motion from the images.

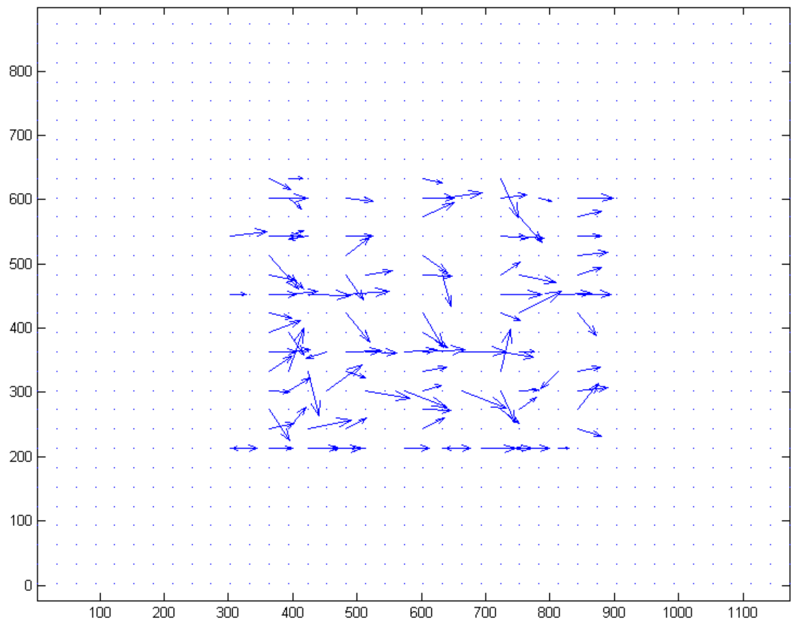


Figure 2. The estimated optical flow indicated by the arrows.

2. Real Data In this image sequence, two people are moving in different directions. The first person in the left part of the image is moving from right to left. The second person in the right part of the image is moving backward to the camera. From our observation, the first person moves fast than the second person. In Figure 3, we can see the estimated optical flows. The estimated optical flows for the second person is better than those for the first person. Further, we can that the motion of the first person is bigger than the second one.

8 Edge Detection By Optical Flow

According to the optical flow constraint equation,

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v = -\frac{\partial I}{\partial t}$$

If $v = 0$, then we can get the following equation,

$$\frac{\partial I}{\partial x}u = -\frac{\partial I}{\partial t}$$

Further we can get

$$\frac{\partial I}{\partial x} = \frac{-\frac{\partial I}{\partial t}}{u}$$

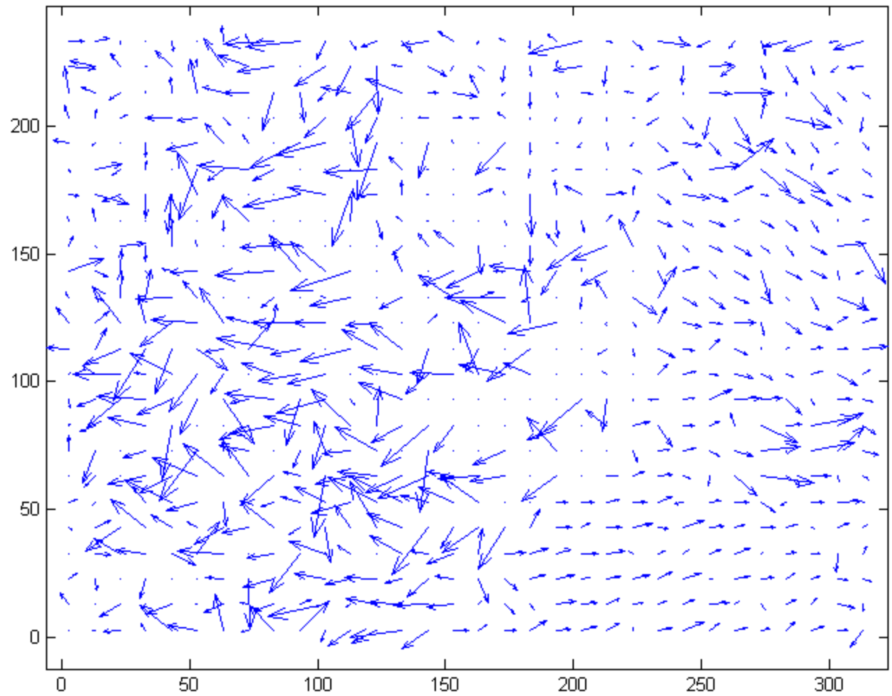


Figure 3. The estimated optical flow indicated by the arrows.

If $u = 0$, then we can get the following equation,

$$\frac{\partial I}{\partial y} v = -\frac{\partial I}{\partial t}$$

Further we can get

$$\frac{\partial I}{\partial y} = \frac{-\frac{\partial I}{\partial t}}{v}$$

Then the image gradients can be estimated according to the (u, v).

In this experiment, we are given two image sequences. The motion for the two sequences are horizontal and vertical translation respectively.

The estimated optical flows of the central image from the image sequence which contains horizontal motion is shown in the Figure 4. We can see that for most of the pixels in the image, the estimated optical flows are same the observed motion direction.

The estimated optical flows of the central image from the image sequence which contains vertical movement is shown in the Figure 5. We can see that for most of the pixels in the image, the estimated optical flows are same the observed motion direction.

After obtaining the optical flows from these two image sequences, we can combine the two OF magnitudes to produce edge strength for each pixel in the central frame. After threholding the edge strength image using different values, we can produce the edge images in figure 6 (b)(c)(d).

In the Figure 6, the correctly detected edge image is shown in (a). Compared with (b)(c)(d), the edge images produced by the optical flow, we can see that the optical flow edge detection method work, but it is not very good.

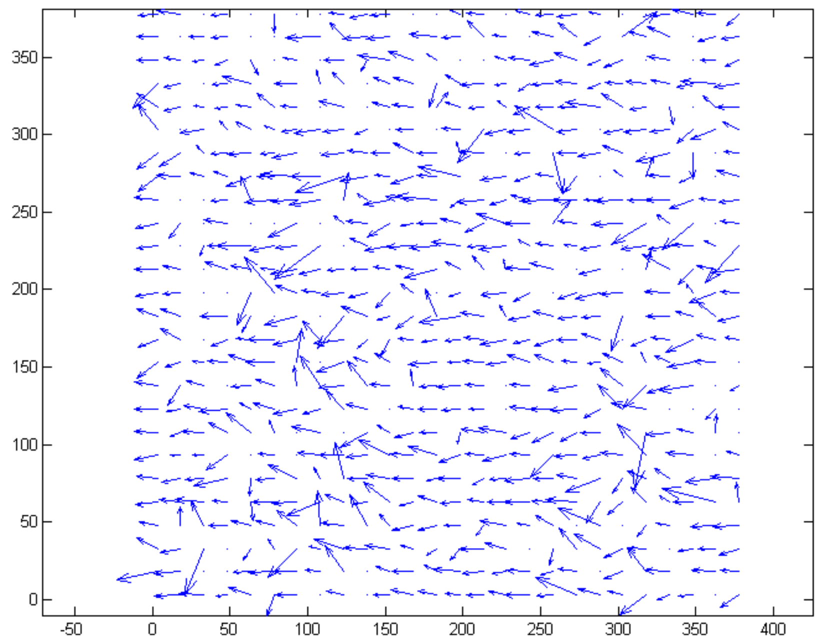


Figure 4. The estimated optical flow indicated by the arrows.

9 Conclusion and Summary

In this project, we implemented the optical flow estimation method introduced in the Computer Vision class. Both synthetic and real image sequences are used to test the implemented method. I found that the optical flow estimation method is very sensitive to the image noise. Further, in order to accurately estimate the motion of the object by optical flow, the texture of the object must have significant image gradients, otherwise, it can not work.

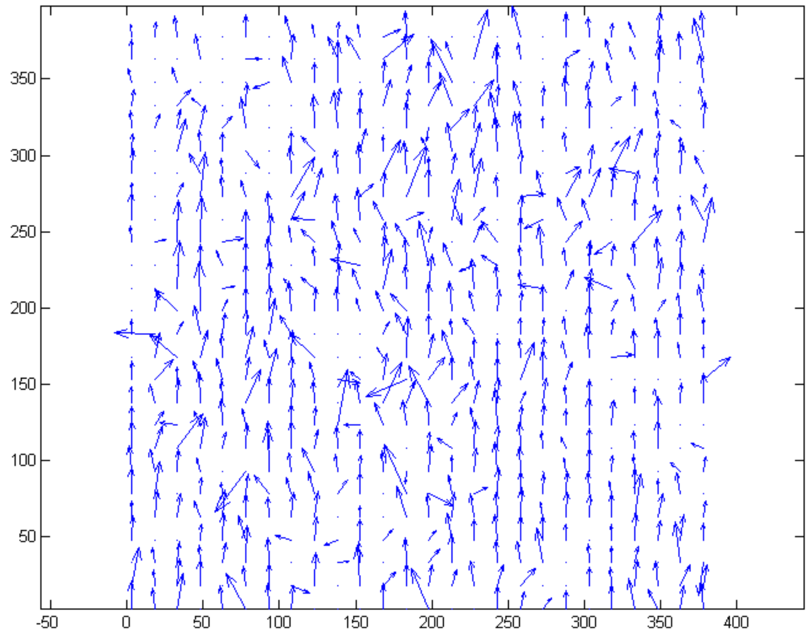
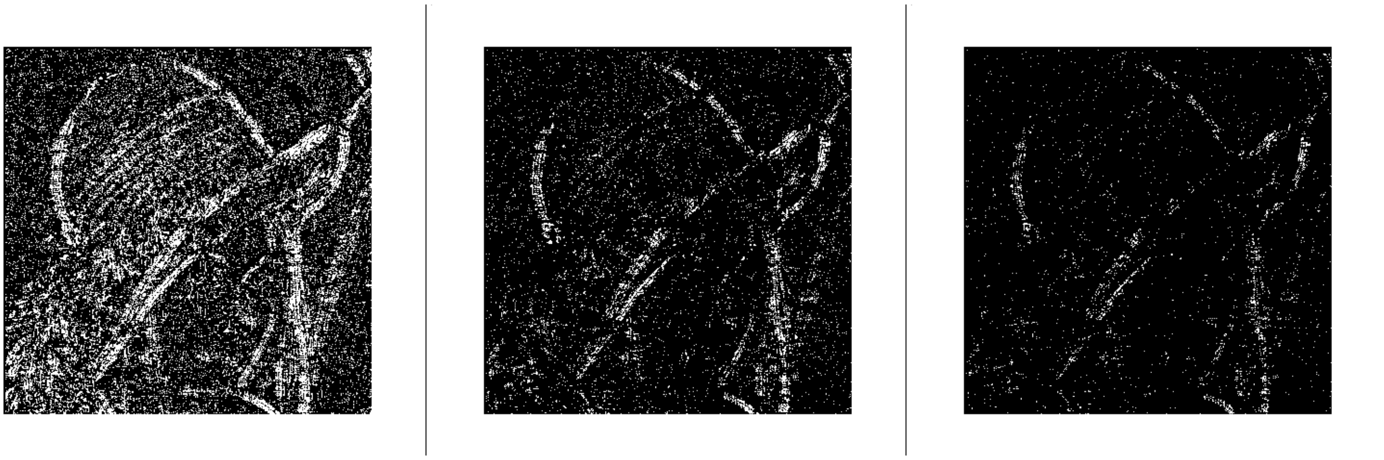


Figure 5. The estimated optical flow indicated by the arrows.

The correctly detected edge image (a)



(b) (c) (d)

Figure 6. (a) The correctly detected edge image and (b)(c)(d) the edge images are threholded by different values.