

# **Computer Vision**

## **Assignment 2 : Secrets of Optical Flow**

### **1 .Basic of Optical Flow**

#### **1.1 Thought Experiments -**

1. **The number-one use of optical flow in visual effects is for retiming a sequence — speeding it up or slowing it down. Describe how optical flow could be used to create slow motion video. You can find the answer in Amazing Slow-Motion Videos with Optical Flow video on YouTube.**

**Answer –**

Interpolation is a method by which we can estimate new frames between 2 consecutive frames. Optical Flow uses the concept to interpolation for estimating the most probable frame which can be fitted between the two frames. For example, let say we have 2 frames, and we need to add one more frame in between two frames, we can simply find the mean of both frames and fit the resultant frames between two frames. This additional frame will convert our videos in slow motion.

2. **In The Matrix, one of the most remembered, iconic use of the tactic comes during the rooftop scene where Neo effortlessly dodges one bullet after another. (Re-)watch “Bullet Time” here and explain briefly how optical flow is used. You may also find this amazing video on bullet-time interesting.**

**Answer –**

Slow motion videos can be created using either recording the video at very high fps rate or by adding frames between existing frames using interpolation. We can deploy multiple number of cameras to catch a greater number of frames per shot and we can combine the frames taken by multiple cameras. Also, we will have camera at different such that we will be able cover 360 degrees spectrum of the shot. Also, we can use the concept of interpolation to find the intermediate frames and add those frames to make video smoother and slower. In the case, they have combined the concept amazingly well to create the illumination. Bullet time is created in The Matrix on a green screen set with a succession of stationary cameras encircling the subject. The cameras are turned on in quick succession, and the generated frames are presented with additional CGI interpolated frames, giving the impression that the camera is moving quicker than the subject.

3. **So breathtaking, heartbreaking and brimming with emotion, WDMC is a journey into the afterlife and deals with a dead man’s attempt to reunite with his wife. Catch**

**a glimpse of the "Painted World" here! You will now describe briefly on how optical flow is used to create this "painterly effect".**

**Answer –**

To get the painterly effect, we will first determine the change vector for each pixel. We must find the optical flow between two frames and then plot this optical flow on the frame. Optical flow works by generating a motion vector map for every pixel in a frame. For each frame, a sub-pixel positional value of  $x$  and  $y$  is created, which can then be compared to the next. This effectively tracks the movement of each and every piece within a composition.

**4. Consider a Lambertian ball that is: (i) rotating about its axis in 3D under constant illumination and (ii) stationary and a moving light source. What does the 2D motion field and optical flow look like in both cases.**

**Answer -**

- In 1<sup>st</sup> case when the Lambertian ball is rotating about its axis and the light source is fixed, we will not be able to find any optical flow even though the ball is continuously moving. This happens because the surface of the ball is even throughout and there is no change in intensity at time  $t$  and  $t+dt$ .
- In 2<sup>nd</sup> case when the Lambertian ball is fixed and the light source is moving, we will get the optical flow even though the ball is not moving. This is because the light source is moving so the shadow and the pixels which are the facing the light will change constantly. So, there will be some difference in intensities of pixel at time  $t$  and time  $t+dt$  which will give us the optical flow.

In both cases, the optical flow is not equal to the motion field.

## **Concept Review**

- 1. List down the important assumptions made in optical flow estimation. Describe each one of them in one-two lines.**

**Answer -**

The two important assumptions made in optical flow estimation are as follows:

- **The observed brightness of any object point is constant over time:** The assumption is based on the consideration that pixels in a small window usually follow the similar displacement and hence the movement of the pixels under considerations in consecutive frames over a window will give the similar optical flow.

- **I Nearby points in the image plane move in a similar manner (velocity smoothness constraint):** The assumption is based on the consideration that pixels in a small window have a similar flow and tend to follow it. Hence, we get similar intensity values over the small window of pixels.

**2. Formalize the objective function of the classical optical flow problem. Clearly mark the data term and the spatial term. What does the objective function imply about the noise distribution?**

**Answer -**

In Optical Flow problem, we make basic assumptions that the intensity of pixel at time  $t$  is equal to intensity at time  $t+dt$ .

$$I(x, y, t) = I(x + u, y + v, t + dt)$$

$$I(x, y, t) = I(x, y, t + dt) + I_x \cdot u + I_y \cdot v$$

$$I_x \cdot u + I_y \cdot v + I_t = 0$$

Here,  $I_x = \frac{dI}{dx}$  (Partial differential of  $I$  w.r.t to  $x$ )

$I_y = \frac{dI}{dy}$  (Partial differential of  $I$  w.r.t to  $y$ )

$I_t = \frac{dI}{dt}$  (Partial differential of  $I$  w.r.t to  $t$ )

The  $I_x$  and  $I_y$  terms in the equation are called data terms and  $I_t$  is the spatial term.

Before finding the optical flow, we must apply Gaussian blur to the image to decrease the noise distribution because the objective function does not include a noise distribution term, we must first pre-process the image before applying it.

**3. In optimization, why is the first-order Taylor series approximation done?**

**Answer-**

The brightness constancy assumption in optical flow is that the intensity of pixels in a narrow window remains constant over time. So,

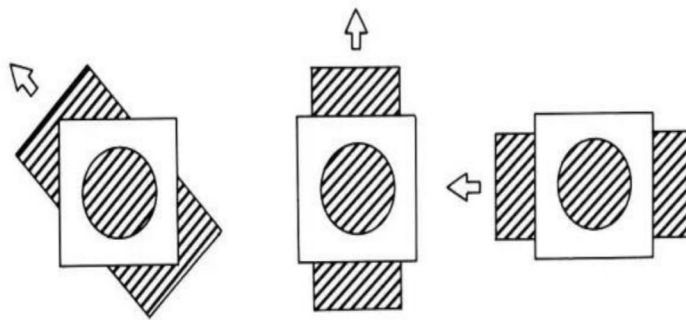
$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

Because the difference is very little, we can use the Taylor series approximation on the given equation. The Taylor series approximation's higher-order terms become very close to 0 and are hence discarded. As a result, we're left with only the first-order Taylor series approximation terms. We just consider the change in intensity rather than the intensity values in each subsequent frame with the help of this optimization.

4. Geometrically show how the optical flow constraint equation is ill-posed. Also, draw the normal flow clearly.

**Answer -**

When we try to get an optical flow for the edges of some patterns then it miscalculates the optical flow. For example, as we can see in the image that with the data in the small circle, we cannot tell which direction it is moving. Due to small size of the aperture the remaining portion of the rectangle given is not visible to aperture. So when we calculate optical flow in such cases there are very high chances that we will calculate it incorrectly.



## **2. Single-Scale Lucas-Kanade Optical Flow**

### **2.3 Analyzing Lucas-Kanade Method**

1. **Why is optical flow only valid in the regions where local structure-tensor ATA has a rank 2? What role does the threshold ( $\tau$ ) play here?**

**Answer -**

We need 2 tensors because we need 2 eigen values of the matrix. The eigen values are directly proportional to the change in x and y directions. The smaller eigen value should be greater than threshold for the pixel to be considered. We the rank of the ATA matrix is one than we will only be able to identify the change in any one of the directions not in both x and y and optical can be in any or both directions. Hence, optical flow is only valid in the regions where local structure-tensor ATA has a rank 2. To identify the best features, we need threshold. Threshold basically checks if the value at certain pixel is high to be classified as a corner or not. In our case, we have taken threshold of value of 0.01 which means we will take all the pixels for which the value is greater than threshold.

2. In the experiments, did you use any modifications and/or thresholds? Does this algorithm work well for these test images? If not, why?

**Answer -**

There are multiple thresholds I have used like the most important one is tau. All the other threshold which are used is for better visualization of results like in case where we have to draw the quiver plot, I have taken some of the points so that the plot looks better. The algorithm works well but it also detects some of the noisy parts of the image as corner.

3. Try experimenting with different window sizes. What are the trade-offs associated with using a small versus a large window size? Can you explain what's happening?

**Answer -**

On performing the experiment using different sizes it is observed that -

- In the case of large window size, the accuracy of result hampers in case of small motion when there is very small change between frames but when there is large motion change then large window size works better than small window size.
- In the case of small window size, we get more accurate results on a pixel level when the motion between two consecutive frames is very less, but the drawback of small size window is we are not able to identify large motion change.

4. There are locations in an image where Lucas–Kanade optical flow will fail, regardless of choice of window size and sigma for Gaussian smoothing. Describe two such situations. You are not required to demonstrate them.

**Answer-**

These are two conditions where Lucas-Kanade optical flow will fail -

1. Speed - When the instantaneous velocity of the point is too large with respect to the window where motion is being considered. Hence, It loses its meaning of temporal derivative.

2. Rotations - The motion around the point has a strong rotational component and thus, even locally, the assumption regarding the similarity of velocities falls.

5. Did you observe that the ground truth visualizations are in HSV color space? Try to reason it.

**Answer -**

HSV stands for Hue, Saturation and Value. Hue in HSV determines the color portion of the image, hue lies between 0 to 360 degrees where each 60-degree bracket has a color assigned to it. Saturation describes amount to gray in particular color it lies between 0 to

1 or 0 to 100% where 0 means completely gray and 1 means completely primary color. The third component which is Value is the amount of brightness or intensity of the color where 0 means completely black and 1 means completely bright. Ground truth visualizations are in HSV color space because optimal flow works on the assumption that the brightness of any pixel is conserved. In HSV, one of the three components is brightness/value. So, HSV color space give better visualization than any of other methods. Hence, HSV color space is used for visualization.