

Assignment 2 The Secrets of Optical Flow

1 Basics of Optical Flow

1.1 Thought Experiments

Q1 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.

Answer: -

To create a slow-motion video, we need to add additional frames in the video which contain information that is not present in the input video. To produce a smooth video, we need to do interpolation of 2 consecutive frames. Interpolation estimates intermediate position of the objects between 2 consecutive frames. To calculate these intermediate positions, we need to get the optical flow of moving pixels and with the help of optical flow, we can create new intermediate frames to add the frames to the original video and slow down the video. Slowing video will work without interpolation as well but the resulting video will not be smooth. The slow-motion video made with interpolation will be smoother.

Q2 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: -

The ‘Bullet Time’ scene in the Matrix is a perfect example of slow-motion video. We can get slow-motion video by 2 methods.

1. Capture more frames for the shot that is to be shown in slow motion
2. Create and add frames using interpolation with help of optical flow.

The 1st method is discussed in detail in the second video of ‘bullet time’ where they have shown the method to create slow-motion videos using many cameras capturing the same scene one after another with very small delay.

The 2nd method can be implemented using interpolation. We need to calculate optical flow of the bullet, then we can add more frames by slowing down the speed of the bullet in the video.

When we combine both methods, we can get slow-motion video as shown in the Matrix movie of ‘Bullet Time’. Several cameras are used to get a 360-degree view of the shot and all the videos from these cameras are slowed down to create the iconic shot of ‘Bullet Time’. While adding additional frames, we

also need to take a video of the actor performing the different actions to dodge the bullet which can easily be done either by superimposing a separate shot of the actor performing the dodge actions or by using extrapolation of frames of the actor by calculating optical flow for the actor in the same shot as the bullet.

Thus, with the help of optical flow and several cameras, we can re-create the bullet time scene.

Q3 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'll now describe briefly on how optical flow is used to create this "painterly effect".

Answer: -

In the painterly effect, the colors are flowing through the scene with help of optical flow. The optical flow of each color is calculated, and the color objects are moved through the frames using the optical flow and keeping the previous frame pixels the same. The painterly effect is created by keeping the colors of the previous frames persistent for a further few frames to make it look like a stroke of paintbrush where colors are spreading which is done using optical flow.

Q4 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 the first case, we will not be able to identify the optical flow due to the lack of good features to detect the optical flow (considering Lambertian ball to be smooth). In this case, we cannot identify the moving pixels due to smoothness hence even though the ball is moving we cannot get the optical flow. We have 2d motion field but no optical flow in this case

In 2nd case if the light source is moving and the Lambertian ball is fixed then we can get the optical flow from the changing pixel values of the shadow of the Lambertian ball. Hence, we can calculate the optical flow. Thus, we get the optical flow of the ball, and it seems like the ball is rotating in the shot even though it is stationary, and the light source is moving. We don't have 2d motion field but there is optical flow in this case.

1.2 Concept Review

Q1 List down the important assumptions made in optical flow estimation. Describe each one of them in one-two lines.

Answer: -

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

- It assumes that the flow is essentially smooth in a local neighborhood of the pixel under consideration – 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
- It assumes that the neighboring pixels under consideration have the same displacement - 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 generally give the similar optical flow.

Q2 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 we assume that the surrounding patch (say 5x5) has constant flow. It means that the intensities of the surrounding pixels in a small window is constant. We can formally write the objective function as

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

Where $I_x = dI/dx$ and $I_y = dI/dy$ which are the partial derivatives of intensities in x and y directions. These are the spatial terms in calculating optical flow

I_t is the temporal term where $I_t = dI/dt$ which is the partial derivative of intensity with respect to time.

We need to apply Gaussian blur to reduce the noise distribution in the image before calculating optical flow. The objective function does not include noise distribution term in it hence we need to pre-process the image before applying the objective function.

Q3 In optimization, why is the first-order Taylor series approximation done?

Answer: -

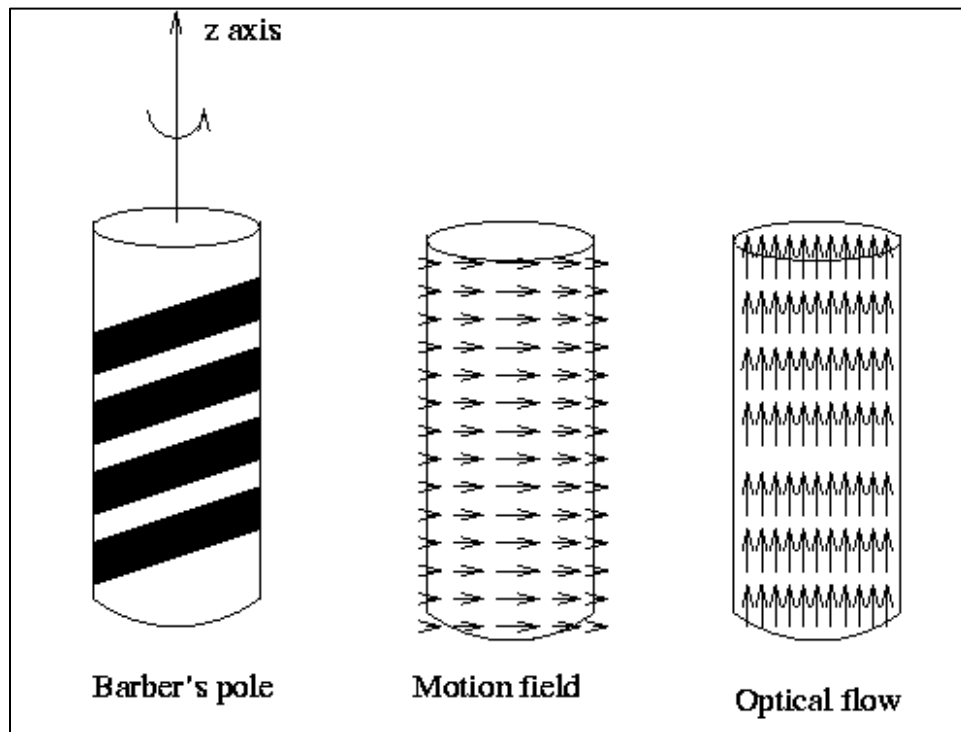
In optical flow we have brightness constancy assumption which states that intensity of the pixels for a particular small window remains constant over time.

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

We can apply the Taylor series approximation on the above equation since the change is very small. The higher-order terms of the Taylor series approximation become very close to 0 hence they are ignored. Thus, we remain with only first-order terms of the Taylor series approximation. With the help of this optimization, we only consider the change in intensity rather than the intensity values in each consecutive frame.

Q4 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. As shown in the above diagram, a Barber's pole moves horizontally, and the motion field is in the X direction, but the calculated optical flow is in the Z direction. This happens because while getting the optical flow we are only concerned with the pixels that move in consecutive frames. We are unable to get the whole picture which might be due to either the Aperture problem or considering only consecutive frames to calculate optical flow.

2 Single-Scale Lucas-Kanade Optical Flow

2.3 Analyzing Lucas-Kanade Method

Q1 Why is optical flow only valid in the regions where local structure-tensor $A^T A$ has a rank 2? What role does the threshold τ play here?

Answer: -

The rank 2 tensors are needed because we need to get 2 eigenvalues of the matrix which are directly proportional to the change in x and y direction. These eigenvalues are similar to the Harris corner detector where the smallest eigenvalue must be greater than some threshold value. If the matrix is rank 1 then the number of linearly independent rows or columns is only 1. This means we can identify the change in only 1 direction which represents an edge. But optical flow on edges is not accurate hence we

need the matrix to have rank 2 so it can have 2 linearly independent rows or columns which help to identify the change in both the directions.

The threshold helps to identify the corners or features which can be easily tracked. The threshold defines the pixels for which optical flow is calculated. The smallest Eigen value of the $A^T A$ matrix must be greater or equal to the threshold to calculate the optical flow. With the help of the threshold, we can go from dense optical flow to calculating only sparse optical flow.

Q2 In the experiments, did you use any modifications and/or thresholds? Does this algorithm work well for these test images? If not, why?

Answer: -

The threshold determines the number of pixels where we can calculate the optical flow. The threshold value is inversely proportional to the number of pixels where we get the optical flow. The algorithm works well for certain test images but misidentifies corners and good features for other test images. This might be due to the presence of noise which misleads the algorithm to think that it has encountered a corner even if there is no corner. We can improve the performance by using Gaussian Smoothing which will reduce the noise for images.

Q3 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: -

In small window size, we get more accurate results on a minute level where there is very little change in the consecutive frames while in large window size, we can identify large motion changes. There is a tradeoff between getting accurate results or identifying the large changes. As the window size increases, the accuracy of the change might decrease but we can identify a larger motion due to the presence of a larger window.

Q4 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: -

The Lucas Kanade fails at following situations

1. Large Movement: When the change in consecutive frames is more than the window size then the algorithm is not able to detect the optical flow for these situations
2. Diagonal Motion: The algorithm is not able to determine the diagonal motion of the edges due to aperture problem. Eg. Barber pole problem for optical flow where we cannot detect motion field of barber pole as shown in 1st part.

Q5 Did you observe that the ground truth visualizations are in HSV colour space? Try to reason it.

Answer: -

In HSV color, V represents the brightness of the image. We have assumed that the brightness remains constant for a small window size in Lucas Kanade. The HSV color space represents these changes more accurately with the help of different brightness levels representing which of the parts are moving in the image. Hence HSV is more suitable to show the optical flow of consecutive frames.