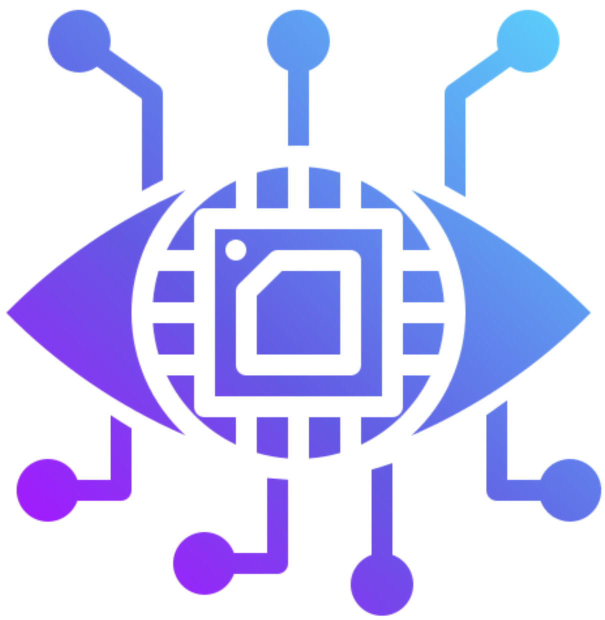
Computer Vision (CS.505)
Assignment 2

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD



Deepti Rawat
2022801016
PhD Scholar

Table of Contents

1 Basics of Optical Flow

1.1. Thought Experiments

1. **Describe how optical flow could be used to create a slow-motion video.**

   For the slow-motion effect, optical flow uses a technique called interpolation.

   Interpolation is a technique with which, given an upper and lower bound measurement of a quantity, one can find an unknown measurement point that lies between them.

   There are two ways to generate a slow-motion video from a regular video.

   - One method is to duplicate the frames i.e., double or triple every frame in succession. For instance, if we have a video with 10 FPS, that means we are getting 10 unique frames in a second. Now, doubling every frame means, now we are getting only 5 unique frames in a second. This might create an effect of slow-motion but the resultant video would be very choppy with not-so-smooth motion sequence. Moreover, the majority of the frames would not contain any information in this video.
   - Another way of creating a slow-motion video is by interpolation i.e., by taking two successive frames of the original video and creating a new third frame that is the result of interpolation of those two frames. Using this process, we can generate a new image to simulate an intermediate behaviour that was not pre-existing. Now, these new images can be used to fill-in between original images, hence creating a smooth slow-motion effect.

   Ref**:**
   Amazing Slow Motion Videos With Optical Flow | Two Minute Paper…

2. **Explain briefly how optical flow is used in the sequence where Neo effortlessly dodges one bullet after another in “The Matrix”.**

   In that particular scene of the movie, optical flow is used to create a visual effect called Bullet time.

   In this technique, a subject is simultaneously captured via multiple still cameras placed around it. Individual frames captured via these multiple cameras at time “t” can be used to create an illusion of motion. This can be achieved via the interpolation technique used in optical flow, in which, given two close points, a third intermediate point is generated which contains new information. For instance, if I have captured a subject at time ‘t’ from two closely placed cameras resulting in View1 and View2, then via optical flow, I can simulate a new view point that lies between View1 and View2. Similarly, this can be done for successive frames captured via other cameras . Hence, if we visualize all these frames, i.e., the originally captured frames and the intermediate frames that are created via optical flow, then we can get a very smooth motion sequence.

3. **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 do the 2D motion field and optical flow look like in both cases?**

   If a point is moving in a 3D scene, its projection on a 2D image plane creates a motion on the image plane known as the 2D motion field corresponding to that moving point in the scene.

   In a 2D image, to measure this motion field, all we can measure is the motion of brightness patterns or the apparent motion that is referred to as the optical flow. However, optical flow might not always be equal to the motion field of the point. Let’s understand this via the following two scenarios.

I. **The 2D motion field and optical flow for a Lambertian ball that is rotating about its axis in 3D under constant illumination.**

If the camera that is capturing the motion of the lambertian ball is still, then the 2D motion field corresponding to the moving 3D point in the ball definitely exists; however, the optical flow might not exist. This is because optical flow is the measure of the brightness patterns in the image and due to constant illumination and the fact that the camera is still, the images captured by this camera would be identical, i.e., the brightness patterns in the image might not change at all. Hence, in this scenario, optical flow doesn't exist, whereas the 2D motion field exists.

II. **The 2D motion field and optical flow for a Lambertian ball that is stationary and a moving light source.**

Here, the lambertian ball is still i.e., there is no physical motion in the original 3D point on the ball and therefore, its corresponding 2D motion field doesn't exist. However, the light source is moving, that means there will not be constant illumination and we can observe motion/change in the brightness pattern in the images captured by the camera.Thus, here optical flow exists. Hence, in this scenario, optical flow exists, whereas the 2D motion field doesn't exist.

1.2. Concept Review

1. **What does the objective function imply about the noise distribution?**

The objective function is utilized in the Lucas-Kanade optical flow approach to explain the correlation between the brightness of pixels in two successively captured frames. The objective function is based on the brightness constancy assumption, which stipulates that the brightness of an image point remains constant throughout time between two frames taken in quick succession.

The Lucas-Kanade optical flow technique determines the optical flow by minimizing the objective function, which is the sum of the squared differences in the brightness of adjacent pixels in the two frames. This approach is used to identify the optical flow given two consecutive frames.

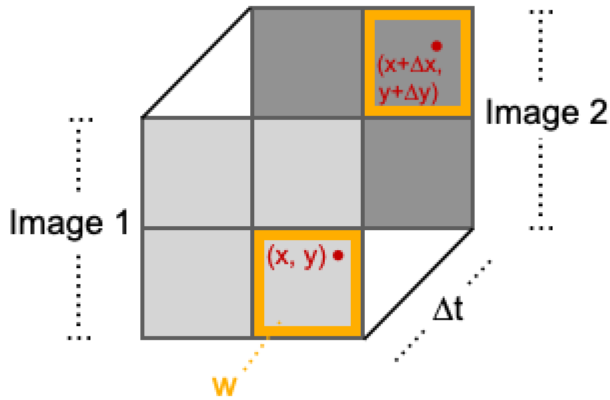
The objective function suggests that the noise in the optical flow estimate is Gaussian, because the sum of squared differences in the brightness of consecutive pixels is proportional to the Gaussian distribution. The quantization error, sensor noise, and fluctuations in pixel brightness brought on by shifting lighting conditions or moving objects are some of the causes of the noise in the optical flow estimate.

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

We know that the motion field is a 2D image projection of velocity of a corresponding 3D point in the scene. This apparent motion of a point on a 2D image plane can be represented via optical flow of the point in the image, where optical flow is denoted by u, v.

To estimate optical flow at each point, is an under-constraint problem. Therefore, we first derive an optical flow constraint equation and then develop an algorithm to estimate optical flow based on this constraint equation. Taylor series approximation allows us to achieve a linear approximation to represent optical flow constraint equation.

Optical Flow Constraint Equation Derivation



- Let's consider a scenario - we have two images of a scene captured in quick succession at time t and t + Δt, respectively. Let's try to analyze a point having location (x,y) in a small window w of the first image. Say, at time t + Δt, that point moved to a new location (x+Δx, y+Δy) in the second image. The resulting displacement of the point from the first image to the second image can be denoted as (Δx, Δy).
- Optical flow or the velocity of that point in the x and y direction can be denoted as:

$$u = \frac{\Delta x}{\Delta t}, \qquad v = \frac{\Delta y}{\Delta t}$$

where, u is the x-component and v is the y-component of optical flow of that point.

- We need to solve for u and v. For this, we need to make a couple of assumptions.
- **Assumption 1**: Brightness of an image point remains constant over time.

i.e., the brightness of the point (x, y) in the first image should be constant at the point (x+Δx, y+Δy) in the second image.

Mathematically, $I(x + \Delta x,\ y + \Delta y,\ t + \Delta t) = I(x,\ y,\ t)$ … **(1)**

- **Assumption 2**: Displacement (Δx, Δy) and time step Δt are very small.
- Through Taylor Series Expansion, a function can be expanded as an infinite sum of its derivatives, i.e.,

$$f(x + \Delta x) = f(x) + \frac{df}{dx}\Delta x + \frac{d^2f}{dx^2}\frac{\Delta x^2}{2!} + \dots + \frac{d^nf}{dx^n}\frac{\Delta x^n}{n!} \quad \dots \textbf{(2)}$$

- From our second assumption, we know that ∆x, ∆y and ∆t must be very small. This implies that in the equation 2, the higher order terms becomes extremely small and we can assume that all higher order terms become zero and thus, we get the first order taylor approximation, i.e.,

$$f(x + \Delta x) = f(x) + \frac{df}{dx}\Delta x \quad \dots \textbf{(3)}$$

- For a function of three variables, equation 3 can be re-written as:

$$f(x + \Delta x,\ y + \Delta y,\ t + \Delta t) \approx f(x, y, t) + \frac{df}{dx}\Delta x + \frac{df}{dy}\Delta y + \frac{df}{dt}\Delta t$$

… **(4)**

- Equation 1 can be expressed in terms of equation 4 as:

$$I(x + \Delta x,\ y + \Delta y,\ t + \Delta t) \approx I(x, y, t) + \frac{dI}{dx}\Delta x + \frac{dI}{dy}\Delta y + \frac{dI}{dt}\Delta t$$

or, $I(x + \Delta x,\ y + \Delta y,\ t + \Delta t) \approx I(x, y, t) + I_x\Delta x + I_y\Delta y + I_t\Delta t$

…

(5)

- Subtracting equation 1 and 5, we get,

$I_x\Delta x + I_y\Delta y + I_t\Delta t = 0$

- Dividing the above equation throughout by ∆t and taking the limit $\Delta t \rightarrow 0$:

$I_x\frac{dx}{dt} + I_y\frac{dy}{dt} + I_t = 0$ …

(6)

- Equation 6 can also be written as, $I_x u + I_y v + I_t = 0$ … **(7)**
- Thus, we have a simple linear constraint equation of optical flow.

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

   <u>*Geometrical Interpretation of Optical Flow Constraint Equation*</u>

- Consider a uv-space and the optical flow at a point in this space is given by vector **u** (bold u indicates that it is a vector), as shown in the **figure (i)**.
- For any point (x,y), we want to measure optical flow (u,v); however, all we have from equation 7 is the constraint line and (u,v) can lie anywhere on this line: $I_x u + I_y v + I_t = 0$.
- Due to this, optical flow estimation problem becomes an under-constraint problem, as shown in **figure (ii)**.

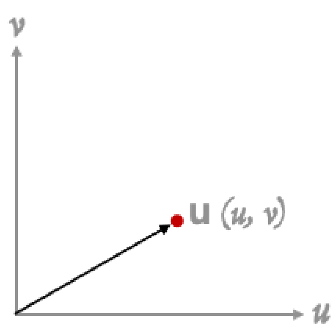


Figure (i)

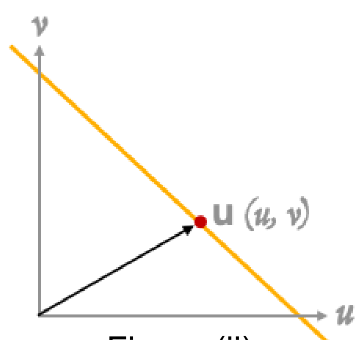


Figure (ii)

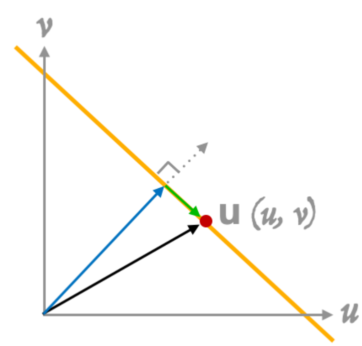


Figure (iii)

- Let's split up optical flow vector **u** into two components, i.e., its normal component and its parallel component. As shown in **figure (iii).** Normal component is normal to the constraint line (blue) and the Parallel component is parallel to the constraint line (green).

$$\mathbf{u} = \mathbf{u}_p + \mathbf{u}_n$$

$\mathbf{u}_n$: Normal Flow

$\mathbf{u}_p$: Parallel Flow

- Normal flow can be computed, given the constrained line in the following way:

  Direction of Normal Flow:

  Unit vector perpendicular to the constraint line:

  $$\widehat{u_n} = \frac{(I_x, I_y)}{\sqrt{I_x^2 + I_y^2}}$$

  Magnitude of Normal Flow:

  Distance of origin from the constraint line:

  $$|u_n| = \frac{|I_t|}{\sqrt{I_x^2 + I_y^2}}$$

  To get Normal flow, multiply the distance and direction of normal flow:

  $$u_n = \frac{|I_t|}{I_x^2 + I_y^2} (I_x,\ I_y)$$

- However, Parallel Flow of optical flow cannot be computed because it could lie anywhere on the constrained line.
- This problem of not being able to determine the parallel flow shows how the optical flow constraint equation is ill-posed.

2 Single-Scale Lucas-Kanade Optical Flow

2.1. Keypoint Selection: Selecting Pixels to Track

Solution in jupyter notebook.

2.2. Forward-Additive Sparse Optical Flow

Solution in jupyter notebook.

2.3. Analysing Lucas-Kanade Method

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

   The local structure tensor, represented by A^T * A, is used to weight the gradients and enforce consistency with the locally observed image intensity.

   The rank of the local structure tensor represents the number of independent gradient measurements in a local neighborhood. If the rank of the local structure tensor is less than 2, it means that there are not enough independent gradient measurements to estimate the optical flow. In this case, the single-scale Lucas-Kanade method will fail to produce accurate results because it cannot estimate the optical flow in regions with insufficient gradient information.

In the single-scale Lucas-Kanade method, τ is used as a threshold to determine whether the eigenvalues of A^T * A are significant enough to be considered in the computation. This way, tau can help us ignore bad regions for optical flow computation. For instance, in a bad region like a low-textured smooth area, both eigenvalues are similar but very small. That means the equation for all pixels in the window are more or less the same and hence, optical flow cannot be computed reliably. In such a case, the value of tau should be such that these pixels are not considered for optical flow computation.

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

Single-scale Lucas-Kanade optical flow method works on two major assumptions. FIrstly, the intensity of pixels in the consecutive images remains constant over time. This assumption is known as the brightness constancy constraint. Secondly, the displacement of the points $(\delta x,\ \delta y)$ from one image to the other image is very small. However, there are certain types of images for which these assumptions are not valid, and as a result, the single-scale Lucas-Kanade method may not perform well. Some examples include:

- Large motions: The single-scale Lucas-Kanade method expects small motions, and its accuracy may deteriorate for large motions. This is because the brightness constancy constraint is not valid for large motions, and the method may not be able to estimate the correct flow in these cases.
- Textureless regions: The single-scale Lucas-Kanade method relies on the gradient information in the image to estimate the optical flow. If the image contains textureless regions or regions with a lot of edges only,

the gradient information may be weak or absent, and the method may not perform well in these regions.

- Motion with changing illumination: The single-scale Lucas-Kanade method assumes that the motions in the image follow brightness constancy assumption, and its accuracy may deteriorate for motions that do not follow it.

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?**

In the single-scale Lucas-Kanade optical flow estimation method, the window size refers to the size of the local neighbourhood used to compute the gradient information and estimate the optical flow. The Lucas-Kanade method uses the assumption that the optical flow in a very small neighbourhood in the scene is the same for all points within the neighbourhood. The choice of window size has an important impact on the accuracy of the optical flow estimation. Using a small window size has the advantage of providing more detailed information about the image intensity in a local neighbourhood. Also, a small window size can assure that the small window assumption is valid. This can result in a more accurate optical flow estimation, especially in regions with strong gradients. However, a small window size is also more sensitive to noise and can lead to deteriorated results in regions with weak gradients or textureless areas.

Whereas, using a large window size provides a more robust optical flow estimation by averaging over a larger region which can reduce the influence of noise and improve the accuracy in regions with weak gradients. However, a large window size can reduce the details of the optical flow in regions with strong gradients or texture.

4. **Did you observe that the ground truth visualisations are in HSV colour space? Try to reason it.**

   HSV space separates the luma, i.e., colour intensity, from chroma i.e., colour information. Therefore, it is widely used in the image processing domain for visualisation purposes. For a similar reason, the optical flow images in the Middlebury .flo file are represented in the HSV colour space.

   The optical flow quiver plots convey the magnitude and the direction of the optical flow in the image. Same information is conveyed in the Middlebury images via HSV space in a visually appealing format that is easy to interpret.

   In the HSV representation, the hue component is used to indicate the direction of optical flow, with different hues corresponding to different flow directions. The saturation component encodes the magnitude of optical flown because saturation represents the intensity of colour and thus higher saturation values indicate larger flow magnitudes. The value component is usually set to 1, which represents the maximum intensity. Therefore, HSV colour space provides a comprehensive way to visualise both the magnitude and direction of optical flow in a single image.