Computer Vision - Assignment 2 Solutions

Laksh Nanwani - 2021701002

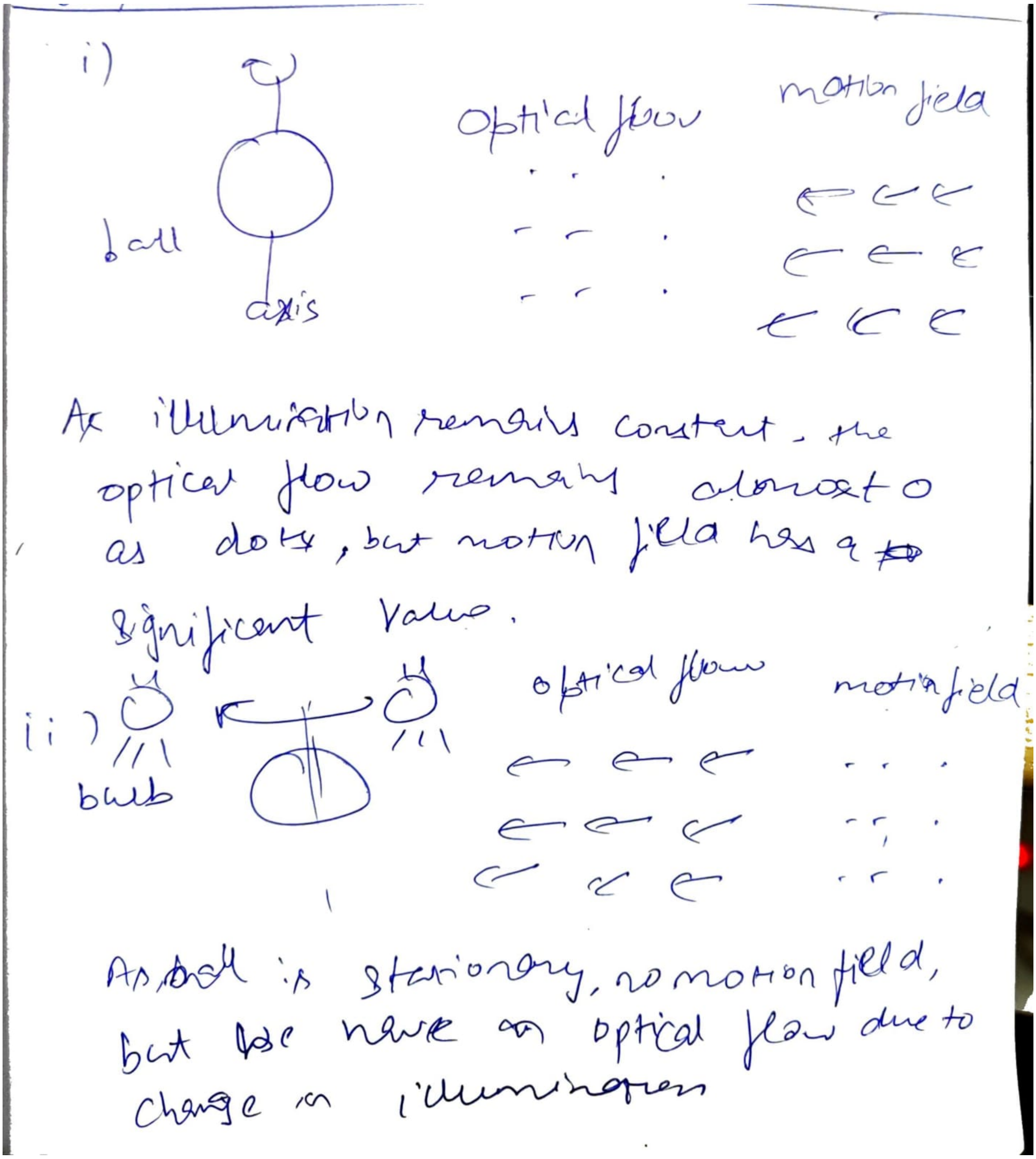
1.1 Thought Experiments

1. A slow-motion video can be made by adding duplicated frames in between the original sequence, but that leaves us with many frames with no new information and hence the video might not be good-looking. Instead of adding duplicated frames, we can use optical flow to estimate an in-between frame between 2 original frames, where the object motion has an intermediate state compared to the original frames. This can be repeated for any 2 continuous frames and hence make a slow-motion video that is smooth and has more details than the 1st method.

2. As mentioned before, optical flow can be used to estimate the intermediate frames and hence intermediate information for making the slow-motion video. In the case of the effect "Bullet Time", when the bullets are fired the video slows down, and hence the need for optical flow arrives. Neo's body movements need to be made smoother between frames and hence the body can be divided into smaller objects such as limbs and optical flow will be applied to these objects to get their motion estimate in the intermediate frames.

3. Painterly effect is an instance of non-photorealistic rendering. We receive a hand-painted look from this effect. To create such effects in videos, Optical flow is used. It follows underlying motion through the scenes and tracks it using brush strokes.
The painterly field is produced via way of means of masking the photograph with paintbrush strokes of random length, radius, and orientations. On the middle of the comb stroke beginning, color is interpolated bilinearly color of the unique photograph. To deliver painterly impact to videos, the motion of the brushstroke among frames wishes to be found. This is completed with the use of optical waft. The gradient-primarily based totally multiresolution approach is used to compute optical waft. An optical waft vector subject is used as a displacement subject to transport the comb stroke to new places in the next frames. Sparse or dense brush strokes can also additionally result. Few brush strokes are deleted in dense areas, primarily based totally on the gap among the comb centers. Delaunay triangulation plus subdividing the mesh into triangles of the vicinity now no longer large than most provided vicinity, and the use of their vertices as brush stroke centers, is completed in sparse areas. So we have the vintage strokes, primarily based totally on movement from the preceding body and new strokes. New ones are uniformly disbursed amongst vintage strokes. The base color is retrieved from the photograph on the stroke middle. The gradient field is decided and is used to calculate every brush`s orientation. This is sustained for the next images.

4. We see that optical flow and 2d motion field do not occur together in both cases. We can see it in the figure below.



1.2 Concept Review

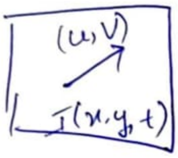
1. Assumptions in optical flow:
    - Brightness Consistency: We assume brightness to be constant for the same points for a small period of time. Taking I to be intensity
      $I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$
    - Small Displacements: Both Spatial and Time displacements are very small.
    - Spatial Coherence Constraint: We assume the motion field around a small area to be constant. We usually choose the window size to be 5x5 or 15x15.

2.

2. A method of estimating apparent motion of scene points in a seq of images is optical flow
This is done for each pixel of the image
But it becomes very difficult to match pixels b/w images

As the time frame is very small, the brightness remains constant.

$$I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$$



To get a linear model, we take taylor series expansion of RHS

$$I(x,y,t) = I(x+u, y+v, t+1) \simeq I(x,y,t) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t}$$

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

$$\Rightarrow \nabla I \begin{bmatrix} u \\ v \end{bmatrix} + I_t = 0$$

To solve this, we use spatial coherence constraint
ie pixels in a small neighbourhood move in same direction..

for a 5x5 window

$$\begin{pmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_5) & I_y(p_5) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_5) \end{bmatrix}$$

$$Ad = b$$

we get constrained soln of this eqn

least square soln is given by:

$$(A^T A) d = A^T b$$

$$d = (A^T A)^{-1} A^T b$$

$$\underbrace{\begin{pmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_y I_x & \sum I_y I_y \end{pmatrix}}_{A^T A} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{d} = -A^T b$$

It's solvable if $A^T A$ is reversible

we ~~get~~ can solve for (u, v)

3. To find linear approximation, we use Taylor series expansion assuming brightness constancy in optical flow constraints and assuming tiny time frames. Also, the motion is assumed to be linear due to very small time gaps. Hence, Taylor series expansion is apt for this application. Also, optical flow is not the original direction of motion of objects but the apparent motion of pixels in the image.

4.

4. As explained in Q2 about brightness consistency.

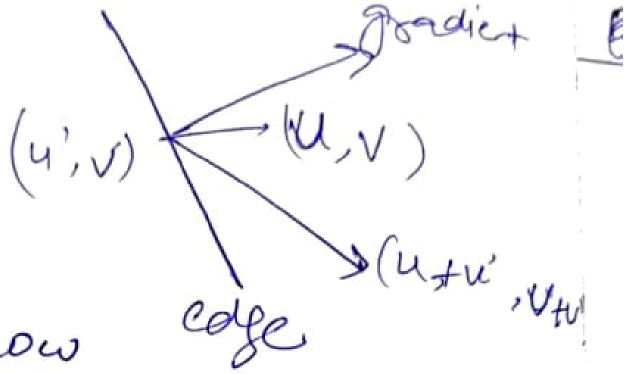
$$I(x, y, t) = I(x+u, y+v, t+1)$$

$$\nabla I \cdot [u', v']^T = 0$$

So,

$I_x u + I_y v + I_t = 0$

is a linear eqn

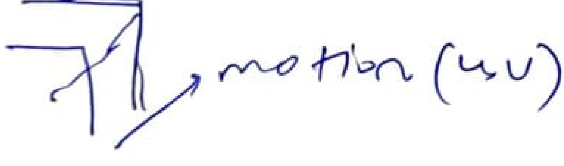


Direction of normal flow

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

$u = u_n + u_p$

This causes apperture problems.



But when looking, ~~wrong~~ a motion looks like Normal flow.

So, locally normal flow can be determined in 2D..

2.3 Analyzing Lucas-Kanade Method

1. We obtain the solution for flow values u,v by taking the inverse of A^TA, and for the inverse to exist the rank of this matrix should be 2. Optical flow is valid at these points because the determinant at these points will be nonzero due to the matrix being square ad full ranked. The threshold removes the flow points with very low eigenvalues, hence removing the effect of noise.

2. We utilize a threshold value to confirm if the points are considered to find good points. The algorithm finds EPE around 2.
We choose corner points using Shi-Tomasi algorithm, after checking the goodness of the points based on the eigenvalues of the matrix A^TA. and we calculate the optical flow at points where both eigenvalues are large.

3. If the window size is small, accuracy is higher. This gives an inverse proportional relationship between the window size and accuracy. But the robustness of the system to varying lighting, size of the image will reduce. Hence, robustness is directly proportional to the window size.
If EPE is increased, more pixels are assumed to have the same optical flow giving incorrect flow calculation. A minimum of two points are required to calculate flow, hence with small window size, the context might get lost.

4. The Lucas-Kanade fails at certain locations in the image regardless of the parameters.
- It will fail when the patch is textureless with only smooth regions making the spatial gradient to be 0. Eigenvalues will reach close to 0 hence won't give reliable optical flow.
- It will fail when there are large motions in the images. After Gaussian smoothing, the number of points for optical flow will reduce by a lot hence not giving reliable optical flow values.

5. Optical flow gives velocity vectors (u, v). The hue is calculated as an angle in the HSV color space. Saturation as the distance from the middle to out gives two cone structures. We set angle as hue, magnitude as saturation so that we map the 3 values from optical flow to HSV values.