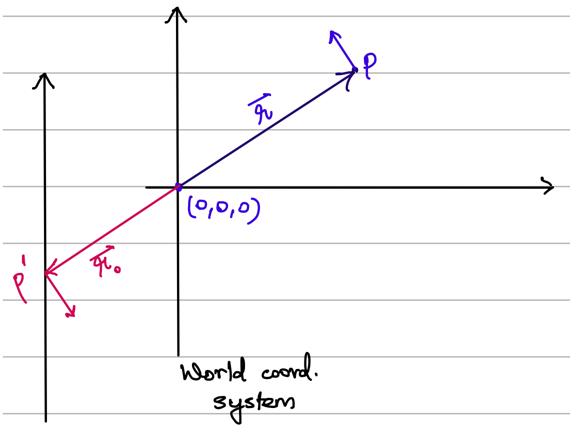
EE702: Lec-10 (12 Feb)

Ego motion

Optical flow ⟶ Visual attention tracking



$$\vec{r}_0 = \frac{f}{(\vec{r}\cdot\hat{z})}\cdot\vec{r}$$

$$\downarrow \frac{d}{dt}$$

$$\boxed{\vec{a}\times\vec{b}\times\vec{c} = \vec{b}(\vec{a}\cdot\vec{c}) - \vec{c}(\vec{a}\cdot\vec{b})}$$

$$\dot{\vec{r}}_0 = \frac{d\vec{r}_0}{dt} = f\frac{\left[(\vec{r}\cdot\hat{z})\frac{d\vec{r}}{dt} - \left(\frac{d\vec{r}}{dt}\cdot\hat{z}\right)\cdot\vec{r}\right]}{(\vec{r}\cdot\hat{z})^2}$$

$$= \frac{f\left(\hat{z}\times\frac{d\vec{r}}{dt}\times\vec{r}\right)}{(\vec{r}\cdot\hat{z})^2}$$

"If something moves in 3D, there must some corresponding 2D plane motion"

Q. Difference between optical flow and motion field

1. Optical Flow:

- **Definition**: Optical flow refers to the apparent motion of pixel intensities in the image due to the movement of objects or the camera itself.
- **Origin**: It is based on the assumption that the intensity of the pixels in the image remains constant as the scene moves.
- **What it represents**: It represents the velocity field of the image, indicating how the pixel positions change over time from one frame to the next, without necessarily distinguishing between the camera motion or the object's motion.
- **Key Concept**: Optical flow is concerned with the movement of **intensity patterns** (or features) across the image plane. It measures how each pixel shifts in the image, but it doesn't directly indicate whether that movement is caused by objects or the camera.
- **Use cases**: Used for tasks like object tracking, motion detection, and camera stabilization.

2. Motion Field:

- **Definition**: The motion field refers to the actual movement of physical points in the scene, typically caused by the motion of objects or the camera.
- **Origin**: The motion field is a theoretical construct that represents the true movement of points in the 3D scene, taking into account the relative motion of objects and the camera.
- **What it represents**: It is a vector field describing how points in the 3D world move over time as seen by the camera. Unlike optical flow, which operates on the 2D image plane, the motion field describes real-world 3D motion.
- **Key Concept**: The motion field represents the motion in 3D space, but it might not always be directly observable from the 2D image. In some situations (e.g., when the camera moves in a way that causes optical flow), the motion field and optical flow can be the same, but they don't always correspond exactly.
- **Use cases**: It's typically used in the context of reconstructing 3D scenes or modeling how objects in the environment are moving.

Key Differences:

- **Representation**: Optical flow is a 2D velocity field (in image space), while the motion field is a 3D velocity field (in world space).
- **Cause**: Optical flow reflects the apparent motion due to the combined effect of both object motion and camera motion, while the motion field represents the real-world motion of objects or points in space.
- **Measurement**: Optical flow is directly measurable from image sequences, whereas the motion field is not directly observable but can be inferred in some cases.

In summary, **optical flow** refers to the 2D movement of pixels in an image due to both object and camera motion, while **motion field** describes the actual movement of points in the 3D world.

Optical Flow

$(x, y \in \text{Image plane})$

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

$(\delta\text{'s very small})$

Taylor...

$$= E(x,y,t) + E_x\,\delta x + E_y\,\delta y + E_t\,\delta t + \cancel{hot} \rightarrow 0$$

$$\Rightarrow E_x\,\delta x + E_y\,\delta y + E_t\,\delta t = 0$$

$\delta t \rightarrow 0$

$$\Rightarrow E_x \frac{dx}{dt} + E_y \frac{dy}{dt} + E_t = 0$$

$$\left(\frac{dx}{dt}, \frac{dy}{dt}\right) \longrightarrow (u,v)$$

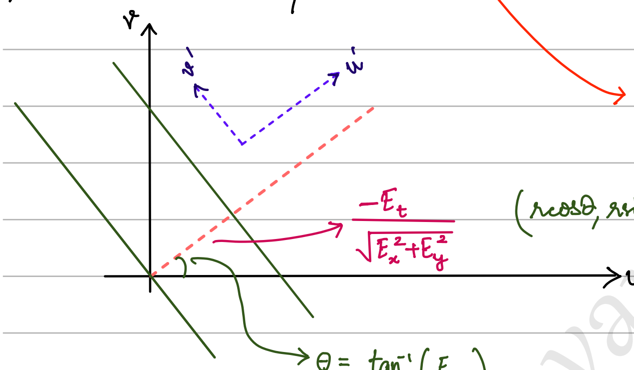
We want to find $\longrightarrow u(x,y),\ v(x,y)$

$$\Rightarrow \boxed{E_x \cdot u + E_y \cdot v + E_t = 0}$$

$\longrightarrow$ Optical flow/Brightness constancy constraint

$\longrightarrow$ We have only one eqn!

So, in the u-v plane:



$\dfrac{-E_t}{\sqrt{E_x^2+E_y^2}}$ $(r\cos\theta, r\sin\theta$ system$)$

$\theta = \tan^{-1}\left(\dfrac{E_y}{E_x}\right)$

"Aperture Effect"

So,

1 eqn, 2 unknowns.

Back to the good old calculus of variation!

Now...

Variational Approach

$$\min_{x,y} \iint \underbrace{\Big[\big[\underbrace{u_x^2 + u_y^2}_{(\nabla^2 u)^2} + \underbrace{v_x^2 + v_y^2}_{(\nabla^2 v)^2} \big]^2 + \lambda \underbrace{\left(E_x \cdot u + E_y \cdot v + E_t\right)}_{\text{Smoothness term}} \Big]}_{F} \, dx\, dy$$

$$\frac{\partial F}{\partial u} - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 0$$

$$\Rightarrow \lambda \left(E_x u + E_y v + E_t\right) E_x - u_{xx} - u_{yy}$$

Sravan K Suresh
22B3936

$$\Rightarrow u_{xx}+u_{yy}=\lambda(E_xu+E_yv+E_t)E_x$$

Similarly, $\Rightarrow v_{xx}+v_{yy}=\lambda(E_xu+E_yv+E_t)E_y$

$$\Rightarrow \boxed{\nabla^2 u=\lambda(E_xu+E_yv+E_t)E_x, \quad \nabla^2 v=\lambda(E_xu+E_yv+E_t)E_y}$$

discretize!

$$\bar{u}_{ij}-u_{ij}=\lambda(E_x u_{ij}+E_y v_{ij}+E_t)E_x, \quad \bar{v}_{ij}-v_{ij}=\lambda(E_x u_{ij}+E_y v_{ij}+E_t)E_y$$

$$\Rightarrow \boxed{\bar{u}_{ij}=u_{ij}(1+\lambda E_x^2)+v_{ij}\lambda E_xE_y+\lambda E_tE_x} \quad \Rightarrow \boxed{\bar{v}_{ij}=v_{ij}(1+\lambda E_y^2)+u_{ij}\lambda E_xE_y+\lambda E_tE_y}$$

$$A_1x+B_1y=C_1$$
$$A_2x+B_2y=C_2$$

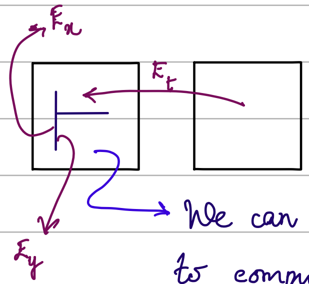
$$\boxed{u_{ij}^{(n+1)}=\bar{u}_{ij}^{(n)}-\frac{E_xu_{ij}^{(n)}+E_yv_{ij}^{(n)}+E_t}{1+\lambda(E_x^2+E_y^2)}\cdot E_x}$$

$$\boxed{v_{ij}^{(n+1)}=\bar{v}_{ij}^{(n)}-\frac{E_xu_{ij}^{(n)}+E_yv_{ij}^{(n)}+E_t}{1+\lambda(E_x^2+E_y^2)}\cdot E_y}$$

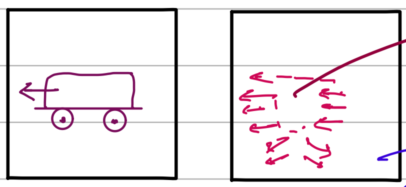
Assuming inter edges/corners/bottom of the camera there is no motion because of the non-ego motion.

Initial condition

(natural boundary condition)



We can take 4 neighbours to compute them to avoid errors due to noise.



$\boxed{E_x, E_y, E_t = 0}$

"needle diagram"

"We can thus detect the motion in discontinuities!!"

Sravan K Suresh
22B3936

Feature-point based Structure-From Motion (SFM)

Ego motion : 3D

Rotation $\longrightarrow R$

Translation $\longrightarrow \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$

$p_1 \quad p_2 \quad p_i \in \mathbb{R}^3$

$p_2 = R\,p + t$

$\to$ (3×3) Rotation matrix

Orthonormal matrix

$(\therefore R^TR = RR^T = I)$

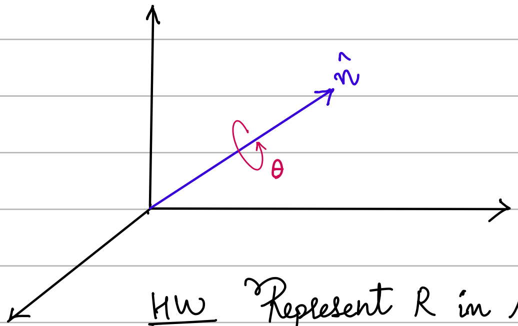
$\det(R) = \pm 1$

(We'll take only +1 because no flipping)

Euler angle $\to$ only for small

Quaternion angle

Rotation



HW Represent R in terms of $(\hat{n}, \theta)$

HW : If $A = K_1 D = D K_2$ where D is symmetric matrix, relate the orthonormal matrices K_1 and K_2

Sravan K Suresh
22B3936