Computer Vision
Spring 2024

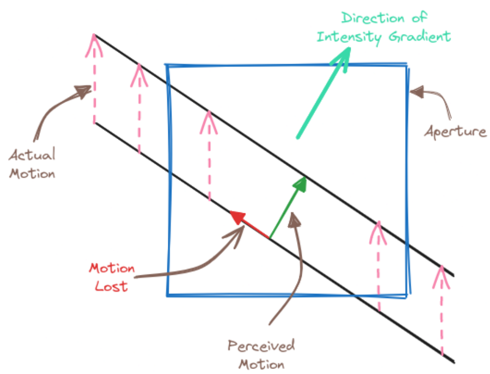
Quiz 2 Solutions

Q1: Aperture Problem

Assigned TA: Mohd Hozaifa Khan

Sample Answer

The aperture problem is the *ambiguity in perceived motion* that arises when estimating motion from a *restricted field of view*. It occurs because we can only measure the component of *motion in the direction of the intensity gradient* and **not along the edge**. As a result, the motion of an edge viewed through a small aperture becomes ambiguous, as the observed motion could result from various possible velocity vectors.



To illustrate, consider a line segment moving across an image. Examining a short portion of the line would not reveal whether it's moving sideways, along its length, or diagonally (figure) – the limited view creates ambiguity. Another example is the Barber-pole illusion.

There are two main approaches to mitigating the aperture problem by imposing additional constraints on Optical Flow Constraint:

- **Lucas Kanade Method** (*local constraint*): A semi-local method which assumes motion is constant within a small $N \times N$ window.
  - Tracking "good" features: track distinctive features like corners or blobs across image frames. These features have a well-defined structure, and their motion can be tracked more reliably, overcoming the aperture problem.
- **Smoothness of Optical Flow Variation** (Horn-Schunck's Method – *global constraint*): This approach assumes spatial smoothness of the velocity field, i.e. optical flow varies smoothly across the image. It is imposed by adding a regularizer term in the brightness constant constraint. Check this for detailed understanding

Rubric

0.5 Pts Ambiguity in motion due to limited field view. Multiple motions lead to the same perceived motion. Cannot detect motion along the edge.

0.5 Pts Solved using any of the following: Lucas Kanade, track distinct features like corners (which give unambiguous motion vectors), global smoothness (HS Method) constraints.

Q2

Assigned TA: Sreenya Chitluri

Sample Answer

To adapt an FCN for depth estimation from monocular images:

- Adjust the output layer to have a single channel for depth.
- Replace classification loss with regression loss.
- Train the modified FCN to do per pixel depth estimation using the data as aquired below.

Data required: Monocular images with corresponding depth maps.

Sources for data: Synthetic scenes, depth sensor data, or specific depth estimation datasets.

Rubric

1. Remodelling architecture -
   - Single output - **0.25**
   - Regression instead of classification - **0.25**
2. Data required - **0.25**
3. Sources of data - **0.25**

Q3

Assigned TA:

Sample Answer

Rubric

Q4

Assigned TA: Sanyam

Sample Answer

Knowns: BRDF, Source Direction Unknowns: Surface Normals
I have allowed "intensity" as part of both knowns and unknowns in the answers because of its ambiguity.
Question was about computing the reflectance map of a surface and not computing the normals given a reflectance map. Most of the students wrote latter which is wrong.

So using iso-contour conics, multiple light sources, least squared error solution, pseudo-inverse method, etc. are wrong answers. 2 different ways:

- Assume a mathematical model for the surface and derive the equations (Ex: Lambertian Model) to calculate reflectance map in terms of knowns and unknowns
- Calibration Method (Take a likewise surface or sphere and position it under same scenario and constraints, get the required reflectance Map for the known object and use it to get the reflectance map for the unknown surface)

Rubric

- Knowns and Unknowns in a Reflectance Map: 0.5 Marks if written both else 0
- Two different ways to compute the reflectance map of a surface: 0.5 Marks if written both else 0

Q5

Assigned TA: Shreya

Sample Answer

The Frankot Chellappa method for estimating shape from surface normals minimizes the error between surface gradients $(p(x.y), q(x, y))$ and the surface gradients of the estimated surface $(z(x, y))$.
The error measure is given by

$$D = \iint \left(\frac{\partial z}{\partial x} + p\right)^2 + \left(\frac{\partial z}{\partial y} + q\right)^2 dxdy$$

Solving using the Frankot Chellappa algorithm:
We minimize the objective function in the Fourier domain. Let $Z(u, v)$, $P(u, v)$ and $Q(u, v)$ be the Fourier transforms of $z(x, y)$, $p(x, y)$ and $q(x, y)$ respectively. i.e.

$$z(x, y) = \iint_{-\infty}^{\infty} Z(u, v)e^{2\pi i(ux+vy)}dudv$$
$$p(x, y) = \iint_{-\infty}^{\infty} P(u, v)e^{2\pi i(ux+vy)}dudv$$
$$q(x, y) = \iint_{-\infty}^{\infty} Q(u, v)e^{2\pi i(ux+vy)}dudv$$

We substitute these in the expression for D, and solve for $Z(u, v)$ that minimizes D using $\frac{\partial D}{\partial Z} = 0$. This give the solution:

$$\tilde{Z}(u, v) = \frac{iuP(u, v) + ivQ(u, v)}{u^2 + v^2}$$

This is the Fourier transform of the surface that minimizes the least square error. The surface itself can be obtained by computing the inverse Fourier transform of $\tilde{Z}$

Rubric

- 0.5 marks for the correct expression of D.
- 0.5 marks for an explanation of how the least square problem is solved

Q6

Assigned TA:

Sample Answer

Rubric