



CAS CS 585 Image and Video Computing - Spring 2021

Assignment 4

This assignment is due midnight (11:59 PM), Friday, March 26th. **Please scan your written assignment as a single PDF file and use GradeScope to submit.**

Exercise 1: Optical Flow

- (a) Draw an optical flow field that shows an object translation that is parallel to the image plane. The motion is twice as fast horizontally than vertically.
- (b) Draw an optical flow field that has a focus of expansion at the principal point and describes a motion an object towards the camera.
- (c) Describe a scenario where the object is not moving but the optical flow field is not zero.
- (d) The Constant Brightness Assumption (CBA) is used in both the Lucas and Kanade Algorithm and the Horn and Schunk Algorithm. Describe how each algorithm handles the fact that the assumption might be violated.
- (e) Explain what it means that "the validity of the CBA depends on the spatial frequency of an image."
- (f) Explain why optical flow perpendicular to the brightness gradient cannot be computed by giving a mathematical argument that uses the CBA equation.
- (g) Spatial derivatives of the flow vector are computed to impose a smoothness assumption in the Horn and Schunk algorithm. 1) Why is smoothness useful? 2) Give the equations for the discrete approximation of each spatial derivative.

Exercise 2: Binocular Stereo (3/23 Lecture)

Suppose you set up the cameras of a binocular stereo system so that the optical axes are parallel and pointing in the same direction and the distance between the centers of projection of the cameras is 20 cm. Both cameras have a focal length of 50 mm and a pixel width and height of 13 micrometers.

How far away is an object imaged for which you measured a disparity of 12 pixels? (Hints: Sketch the geometry. Make sure not to drop the physical units in your calculations.)

Exercise 3: Deep Learning and Computer Vision

- (a) Explain briefly but concisely what it means that a neural network layer performs a convolution of filter size $5 \times 5 \times 3$ and stride 2.
- (b) What is the difference between instance segmentation and semantic segmentation? (1 sentence)
- (c) What does "max unpooling" mean?
- (d) What is meant by a "one hot" representation? What is a word vector representation? Which representation yields better results with deep networks that address the task of automatically developing captions for images to help the visually impaired?

[Margrit Betke, Professor](#)
[Computer Science Department](#)
Email: betke@cs.bu.edu