

CS 223b: Introduction to Computer Vision

Assignment 4: Stereo Vision

Due date: **Monday, February 9th 23:59 PST**

You may work in **teams of up to 3 persons**

Submission via email with subject "Assignment 4: [NAMES]" to cs223b09@gmail.com, where [NAMES] is a comma separated list of the full names of everyone in your group.

1 Stereo Vision (25 pts)

For this assignment we ask that you use OpenCV for the code portions.

1.1 Three Camera Stereo Geometry (20 Points)

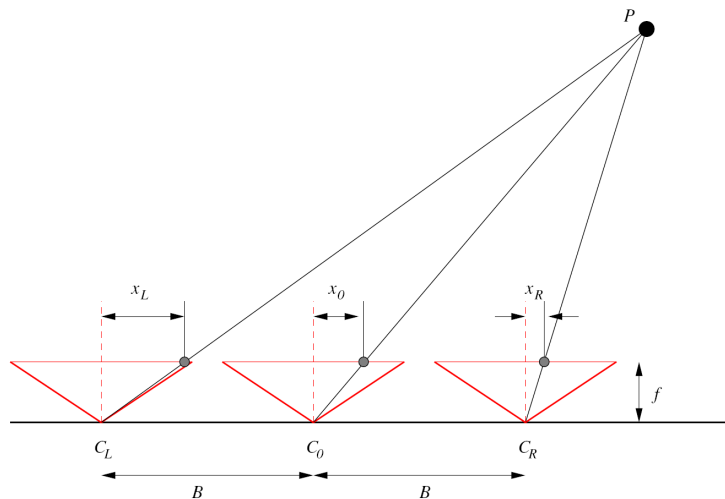


Figure 1: A diagram of the three camera setup used in this assignment

The above diagram shows one possible layout for capturing 3-camera stereo. The distance between consecutive cameras is B . All three cameras are identical: they have the same focal length f and their image planes are parallel to the line of camera centers C_L, C_0, C_R . A point P is imaged at positions x_L, x_0, x_R in the left, central and right camera respectively.

1.1.1 Disparities

Consider the three-camera setup shown in Fig. 1. The disparity of P in the left camera (with respect to the central camera) is given by

$$d_L = x_L - x_0$$

Similarly, the disparity of P in the right camera (with respect to the central camera) is given by

$$d_R = x_R - x_0$$

Show that $d_L = -d_R$.

1.1.2 3D Reconstruction

You are required to write code which given three images from cameras in the layout given in the diagram, computes the disparity of each pixel in the central camera. Consider a pixel with coordinates (x_0, y_0) in the central camera. Suppose it has disparity d_0 in the left camera. Thus, it corresponds to the pixel $(x_0 + d_0, y_0)$ in the left camera and $(x_0 - d_0, y_0)$ in the right camera. Our job is to estimate d_0 for each pixel in the central camera image.

Let I_L, I_0, I_R be the images from the left, central and right cameras respectively. Assume that the disparity of each pixel is an integer and lies in a range $[D_{min}, D_{max}]$. We will assume that corresponding pixels in different cameras have the same color. (This is a reasonable assumption when the objects in the scene are diffuse and the cameras are identical. This can be achieved by using a single camera which moves across three positions). Now suppose we have hypothesized that a pixel (x, y) has disparity d . How can we evaluate the quality of our hypothesis? We shall describe two ways, which you are required to implement.

Sum of Absolute Differences

We can use what is known as the sum-of-absolute distances (SAD) approach. For a pixel (x, y) in the central image and a hypothesized disparity d , we can define:

$$SAD(x, y, d) = \sum_{|x'|\leq W, |y'|\leq W} |I_0(x + x', y + y') - I_L(x + d + x', y + y')| + \sum_{|x'|\leq W, |y'|\leq W} |I_0(x + x', y + y') - I_R(x - d + x', y + y')|$$

This computes the sum of absolute color differences between a window of pixels around (x, y) in the central image I_0 and the corresponding windows in the left and right images. The size of the window used is $2W + 1 \times 2W + 1$. The idea is that if the disparity is correct, the windows are similar and $SAD(x, y, d)$ is small but if the disparity is not correct then $SAD(x, y, d)$ could be large since we are taking color differences between independent pixels. We can estimate the disparity as:

$$Disparity(x, y) = \operatorname{argmin}_d SAD(x, y, d)$$

That is, for each point we choose the disparity which minimizes the sum of absolute differences. **Important Note:** For color images, the difference of two color values is defined to be the sum of the absolute values of the differences in each color channel. If I, I_0 are two color images then

$$|I(x, y) - I'(x', y')| = |I_{Red}(x, y) - I'_{Red}(x', y')| + |I_{Green}(x, y) - I'_{Green}(x', y')| + |I_{Blue}(x, y) - I'_{Blue}(x', y')|$$

Median of Absolute Differences

Instead of adding all pixel color differences over a window, we could take the median difference. Specifically, we define

$$MAD(x, y, d) = \operatorname{Median}_{|x'|\leq W, |y'|\leq W} |I_0(x + x', y + y') - I_L(x + d + x', y + y')| + \operatorname{Median}_{|x'|\leq W, |y'|\leq W} |I_0(x + x', y + y') - I_R(x - d + x', y + y')|$$

and can estimate the disparity as:

$$Disparity(x, y) = \operatorname{argmin}_d MAD(x, y, d)$$

Even at the correct disparity all pixels in a window may not match across images, such as due to occlusion or depth discontinuities. Taking the median is an attempt to be robust to such a scenario.

Implementation Details

We ask that you not use any of the built in OpenCV stereo functions. OpenCV should only be used for loading and saving images.

You can ignore the disparity near the edges of the image where the window crosses the image boundary. For your output, produce a *disparity map* which is a grayscale image where the intensity of each pixel is equal to 6 times the disparity value estimated for it (there is nothing inherently special about the number 6, it just happens to work for the images we give you). Use the disparity range (d) 0 to 42 pixels. Run your code using *cones-left.png*, *cones-right.png*, and *cones-center.png*. Produce an image using *SAD* and a 5x5 window, and save it as *SADdisparity.png*. Produce an image using *MAD* and a 7x7 window, and save it as *MADdisparity.png*.

1.1.3 Camera Layout

Is the 3-camera layout used in this assignment (Figure 1) the best layout possible? If so explain what advantages it has over other layouts. If not state a better 3-camera layout and explain its advantages over the layout from Figure 1. Your answer need not be more than short paragraph.

1.2 The Fundamental Matrix (5 Points)

In this section we ask you to answer several questions about the fundamental matrix. For each question provide both the answer and an explanation.

1. If F is the fundamental matrix of the camera-pair (P, P') , then what is the fundamental matrix for (P', P) ?
2. If for a point x in the first image, the corresponding epipolar line is Fx , then what is the epipolar line corresponding to x' in the second image?
3. Deduce the left and right null-space of F .

1.3 Submitting

Submission Checklist

- Proof that $d_L = -d_R$ (in writeup)
- OpenCV code for computing disparity using *SAD* and *MAD*.
- *SADdisparity.png* computed with a 5x5 window.

- *MADdisparity.png* computed with a 7x7 window.
- Answer to the camera layout question (in writeup)
- Answer and explanation for each of the three fundamental matrix questions (in writeup)

When finished, submit all sourcecode, all images produced, and a .doc or .pdf writeup. Submission should be via an e-mail to cs223b09@gmail.com with the subject “Assignment 4: [NAMES]”.