CS376: Computer Vision: Assignment 4 Due: Nov. 16th, 11:59 PM

**Instruction:** 100 points in total.

**Submission:** See the end of this document for submission instructions.

# 1 Short answer problems [40 pts]

1. **(10 points).** In stereo matching, what is the search space for each pixel of one image in the other image? What is the meaning of image rectification?

The search space for each pixel in the first image is the set of all pixels in the corresponding epipolar line in the other image. Image rectification is the process using the camera parameters of the source images to “warp” and project the content of them onto new images on a plane parallel to the optical centers of the sources. This makes all horizontal scanlines between them epipolar.

1. **(10 points).** In stereo matching, what is the difference between determining the target pixels indepen- dently and determining the target pixels jointly? Please list three criteria for determining the target pixels jointly. What is the approach for solving the joint optimization procedure?
2. **(10 points).** In k-nearest neighbor classifier, what is impact of varying *k* on the classification accuracy? How to determine the value of *k*?

Increasing k increases the accuracy of the classification, but also increases the amount of computations. Trial-and-error accelerated by machine learning can be used to find the best K-value. Increasing K and measuring the accuracy on the machine learning dataset until the increase in accuracy has diminished to a point that increasing K is not worth the extra computations.

1. **(10 points).** List two differences between non-parametric methods and parametric methods for visual recognition.

# 2 Programming: Stereo Matching (60 points)

The goal of this assignment is to implement a simple window-based stereo matching algorithm for rectified stereo pairs. You will be using the following stereo pair: Following what we have discussed in class, pick a

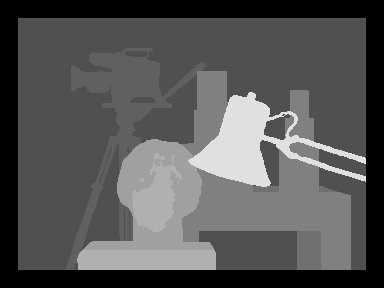


Figure 1: Illustration of the data set. (Left) Source Image. (Middle) Target Image. (Right) Ground-truth disparity map.

window around each pixel in the first (reference) image, and then search the corresponding scanline in the second image for a matching window. The output should be a disparity map with respect to the first view. Note that key parameters should include the search window size, disparity range, and matching function. You should report comparisons for different values of these parameters:

* Search window size : show disparity maps for several window sizes and discuss which window size works the best (or what are the tradeoffs between using different window sizes). How does the running time depend on window size? **Bonus point:** How about combing multiple window sizes?
* Disparity range: what is the range of the scanline in the second image that should be traversed in order to find a match for a given location in the first image? **Bonus point:** How about use the smoothness in disparity ranges to accelerate the computation?
* Matching function: try sum of squared differences (SSD) and normalized correlation. Discuss in your report whether there is any difference between using these two functions, both in terms of quality of the results and in terms of running time.

In addition to showing your results and discussing implementation parameters, discuss the shortcomings of your algorithm. Where do the estimated disparity maps look good, and where do they look bad?

Grading considers the following factors:

* **30 points.** Implement a basic algorithm with a given window size, disparity range, and matching function. Please try sum of squared differences (SSD) and normalized correlation.
* **10 points.** Analyze the effects of the window size. Please display results of three window sizes, one small, one of median size, and one large.

A picture containing chart

Description automatically generated

A picture containing chart

Description automatically generated

Chart

Description automatically generated

The changing the window size results in a tradeoff between two variable determining the quality of the depth map-- noise and precision. Smaller windows mean more precision but increased noise, while larger windows result in less precision but less noise as well. This occurs for the same reason that smoothing filters with larger kernels result in reduced noise but blurrier edges.

* **10 points.** Analyze the effects of changing the disparity range. Show results of using three values for the disparity range.

Chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generatedIncreasing the disparity range seems to have little-to-no drawbacks in terms of the results—there is both increased precision and reduced noise with larger ranges. However, larger ranges means more calculations and therefore a slower runtime. The accuracy/time tradeoff exists largely for the same reasons that increasing k in k nearest classification improves accuracy. There also seems to be diminishing returns after a certain range—which makes sense considering that as you increase the range, the amount of windows left to scan decreases, which means that the chance that the best match is less likely to be in the unscanned range. Additionally, increased ranges means that the relative difference between depths seemingly decreases. Increasing the range from 25 to 40 made the difference between the closest objects and farthest objects seem smaller.

* **10 points.** Compare the effects of using sum of squared differences (SSD) and normalized correlation.

Discerning the differences in results between SSD and normalized correlation is difficult. The results have had unpredictable differences that I simply don’t know the reason for. However there are certain features of the tsukuba image that consistently appear in one but not the other. The most apparent features that appears in normalized correlation but not in SSD is the space between the head and arm of the lamp. However, there are other features, like the box on which the sculpture sits, that the normalized correlation incorrectly gives divots on its sides. I don’t think one is objectively better than the other, I think they have situational uses. Unsure what situations would be better for SSD and what would be better for normalized correlation, however.

# Submission instructions:

Create a single zip file so submit on Canvas that includes

* Your well-commented code, including the files and functions named as specified above.
* A **PDF** writeup of your results with embedded figures where relevant. Please do not include any saved matrices or images etc. within your zip file.