Final Project: Stereo Matching

DIP Teaching Stuff, *Sun Yat-sen University*

Well, your DIP course is drawing to a close. We hope you enjoy this journal and have learned something useful. At last, we are going to develop an interesting application – Stereo Matching. This project will take up a bigger portion of your final score, compared with your mid-term exam. Hence please take it seriously.

1 Introduction

Have you ever wondered how human eyes perceive the world in 3D? In this final project you are about to find out! The principle is actually quite simple: if there are two coins of identical sizes in front of you, the closer one will appear bigger, and the further one will appear smaller.

This observation leads us to the definition of "disparity". Given a pair of images observed by the left eye and the right eye respectively, as shown in fig. 1, the disparity map D_L of I_L is defined as

$$I_L(x, y) = I_R(x - D_L(x, y), y)$$

That is, the disparity of a pixel p in the left image is the difference in x-coordinate between p and its correspondence q in the right image (note that the left and right images here have been pre-processed so that all pixel correspondences lie on the same horizontal lines). Analogously, the disparity map D_R of I_R is defined as

$$I_R(x, y) = I_L(x + D_R(x, y), y)$$

Stereo Matching is the task of estimating the disparity map from pre-processed left- and right-eye images (stereo images) and it is one of the most fundamental problems in computer vision. Fig. 1(c) shows the left disparity map between Fig. 1(a) and 1(b). There are tons of methods proposed for stereo matching and today you are going to implement some of the most populars.

(a) left-eye image I_L (b) right-eye image I_R (c) left-eye disparity map D_L

Figure 1: *The Aloe dataset*

2 Requirement

Here we provide a simple explanation of implementing some local stereo approaches. Given a stereo pair $\{I_L, I_R\}$ and a largest possible disparity value $d_{\max}$, the disparity map D_L is computed by

$$D_L(x,y) = \underset{d\in\{0,1,\ldots,d_{\max}\}}{\arg\min} \ \operatorname{dist}\left(F_L(x,y), F_R(x-d,y)\right)$$

where $F_*(x,y)$ is a feature vector extracted from I_* at location (x,y) (e.g. a 5×5 patch centered at (x,y)), and $\operatorname{dist}(P,Q)$ computes a matching cost between two feature vectors (e.g. sum of the squared color differences of all corresponding pixels in patches P and Q).

2.1 Basic Task

Download and unzip the file ALL-2views.zip. You will find 21 test cases in it. Each test case consists of a left-eye image (view1.png) and a right-eye image (view5.png), accompanied with their "ground truth" disparity maps (disp1.png and disp5.png), which will be used to evaluate your own results. The largest possible disparity for all test cases is 79. Note that the intensities of all ground truth disparity maps have been multiplied by a factor of three for better visualization. Zero values in ground truth are reserved to indicate unknown disparities. Now you are required to

1. Implement a test program to evaluate the quality of your disparity maps. Your program should output the percentage of bad pixels in each of your disparity maps, where bad pixels are those whose errors are greater than one when comparing to ground truths (without multiplication).

2. Implement a local stereo matching algorithm using "Sum of Squared Difference (SSD)" as matching cost. Compute both the left and the right disparity maps for all test cases, and visualize them by scaling their intensities by a factor of three. Save each disparity map to a file named testcasename_disp1_SSD.png or testcasename_disp5_SSD.png. Also write down the matching cost function and the quality of your disparity maps.

3. Implement a local stereo matching algorithm using "Normalized Cross Correlation (NCC)" as matching cost. Visualize and save your estimated disparity maps to testcasename_disp1_NCC.png or testcasename_disp5_NCC.png. You are required to figure out the meaning of NCC by yourselves and write down the formula of the NCC matching cost. Also write down the quality of your disparity maps.

4. Add a small constant amount of intensity (e.g. 10) to all right eye images, and re-run the above two methods. Analyze how the intensity change affects the results (i.e. the quality) of the two methods. Explain in which ways that NCC is a better matching cost than SSD.

5. Cost aggregation using Adaptive Support Weight (ASW) is a powerful local stereo approach. You are required to implement this approach described in the CVPR'05 paper "Locally Adaptive Support-Weight Approach for Visual Correspondence Search" (to save time you are recommended to read mainly sec. 2.6 and sec. 3 of the paper). Visualize and save your estimated disparity maps to testcasename_disp1_ASW.png or testcasename_disp5_ASW.png. Also write down the quality of your disparity maps. Note that this algorithm is in general a bit slow.

6. Explain the idea of the ASW paper in question 5. Also explain in which ways ASW is better than NCC. Why?

2.2 Advanced Task (Optional)

In this advanced task, you are going to go beyond the local stereo matching methods.

1. Implement the Semi-Global stereo matching (SGM) method described in the CVPR'05 paper "Accurate and Efcient Stereo Processing by Semi-Global Matching and Mutual Information". We recommend you to mainly focus on sec. 3.3 of this paper. Please skip sec. 3.2 and use a simple matching cost instead (e.g. sum of absolute differences on 3×3 patches). In the paper the matching cost is aggregated along 16 paths. In order to simplify your implementation, you could use as less as four paths (along horizontal and vertical directions). Visualize and save your estimated disparity maps to testcasename_disp1_SGM.png or testcasename_disp5_SGM.png. Also write down the quality of your disparity maps.

2. Explain the main differences between local stereo approaches and SGM.

2.3 Language

You can select any language as you like.

2.4 Submission

1. Well-documented codes.

2. A project report, at least including 1) your answers to all questions mentioned above, 2) your implementation details, and 3) discussions of the disadvantages of the paper(s), and your improvements / potential solutions.

2.5 Plagiarism

You can discuss this project with your classmates. However, collaborating with anyone else during this individual project is strictly prohibited. This includes but not limits to sharing the problem solutions with others or plagiarizing materials from the Internet. If we believe that you have undermined the integrity of this project, we reserve the right to disqualify you from this course. Use your best judgment.