

## EE 702: Assignment 2

### Stereo Vision

#### Introduction

The depth is estimated using three different measures of similarity between the patches and two different methods. Multi ocular and Bino ocular both are implemented

##### Similarity measures:

1. SSD (Sum of Squared Difference)
2. SAD (Sum of Absolute Difference)
3. NCC (Normalised Cross Correlation)

The two different methods used are as follows

##### 1. Full image depth

Here the similarity between the patches is found using the given input images directly (without finding edge features)

##### 2. Edged depth

Here edges in the image are found. The depth is calculated only at the edges. Later this depth is spread into the region using by solving iterative optimization problem which also takes regularization into account.

##### No of input stereo images:

1. Multi Ocular (4)
2. Bino Ocular (2)

#### Code:

Run 'main\_entire.m' to generate and save results in respective folders.

'stereo2.m' : calculates depth for full image depth method

'stereo.m' : calculates depth using edged depth method

Other scripts are written to save results in respective folders.

#### Results:

1. results/edging : contains bino ocular and multi ocular depth images calculated using edged depth method.
2. results/full\_image: contains bino ocular and multi ocular depth images calculated using full image depth method.

#### Analysis:

Different similarity measures:

##### 1. SAD

SAD is the sum of absolute differences between the patches. The results obtained using this methods are observed to be affected by intensity spikes in the input images. It allows user to detect directionality in the images effectively. Therefore black regions (undesigned depth pixels formed due to occlusion) are not observed in the case of SAD. But being a nonanalytic function may affect the optimization results in case of edged depth method, where we intend to spread the depth using iterative approach.

## **2. SSD**

SSD is the sum of squared differences between the patches. The results contain few dark patches. This can be incorporated to partial occlusion of objects while camera is moved to capture next image (NOTE: We are using fixed rectangular window. This occlusion can be addressed by using variable shape window, whose shape will be determined by the availability of common points in the images i.e. non occluded portion of the scene.)

## **3. NCC**

NCC is supposed to give the best results amongst all. It is computationally more expensive than SSD and SAD. The problem of occlusion is present here also. But it effectively addresses the issues of high intensity spikes. But NCC fails to detect depth discontinuities (object boundaries) as effectively as SAD, because since it is normalised correlation, it tends to smoothen things too much. But overall on other aspects it performs better.

### **Multi ocular:**

Multi ocular considers multiple input images (4 in this case) taken from different locations (here we assume them to be taken from points on a straight line with a fixed distance spacing). Then pairwise disparity is calculated for adjacent camera locations. The final output is the weighted sum of the pairwise disparities where weights depend upon the distance between locations of capture.

Multi ocular provides larger field of view and hence reduces error caused due to occlusion and short sighting which are generated due dependency on neighbourhood pixels.

The two different methods used are as follows

### **1. Full image depth**

It is computationally more expensive but gives better results. Its superiority over other is highly enhanced in case of multi ocular depth estimation. Better results are obtained because we considering every point in the window for similarity measures and assigning depths (which is obviously better than considering only edges or any other feature points for that matter ).

### **2. Edged depth**

Here edges in the image are found. The depth is calculated only at the edges. Later this depth is spread into the region using by solving iterative optimization problem which also takes regularization into account. Computationally it is cheaper, but results obtained are not as good as that of Full Image depth, especially in case multi ocular.