University at Buffalo, SUNY

CSE 573
Fall 2016
Final Project
16th December 2016

Disparity for Stereo Vision – Block Matching and Dynamic Programming

Submitted by,

Arun Sharma, 50206920

Jayaraj Sajjanar, 50208475

Introduction

Stereo Vision is a technique in which we find the depth of the image by finding it's disparity map of the images. Given two images we can infer depth by matching pixels form the current image to the corresponding pixels in the reference (shifted) image. The change in the shift of the pixels gives the disparity, i.e. larger the shift, brighter the pixel. From this we can identify the foreground and background objects. The following is the diagram for finding the disparity map.

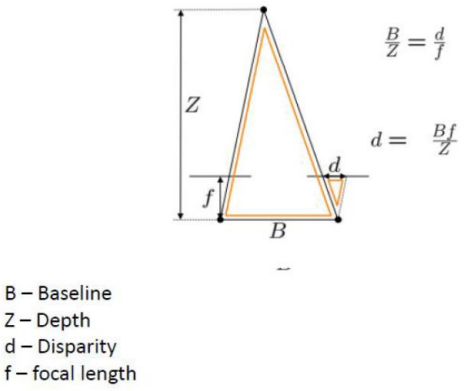


Fig.1 Disparity Map[3]

Overview of the project

In this project, we implemented block based matching and dynamic programming approach techniques for finding disparity map of two images. For given disparity map we calculated their respective MSE value and analyzed the results tuning different parameters.

Your Approach (Overall System)

Block Based Algorithm

A block based algorithm matches the macroblocks from current frame with the referenced frame. This technique can be divided mainly into three components

- Block Determination
- Search Method
- Matching Criteria

The first component is block determination initialize the position and size of the blocks in the current frame and the starting location in the reference image.

The second component is the search method which specifies where to look in the reference frame in which an exhaustive search is matching every block with the reference frame.

The third component is the matching criteria where we set a certain criterion for the matching the block between the current frame and the referenced frame.

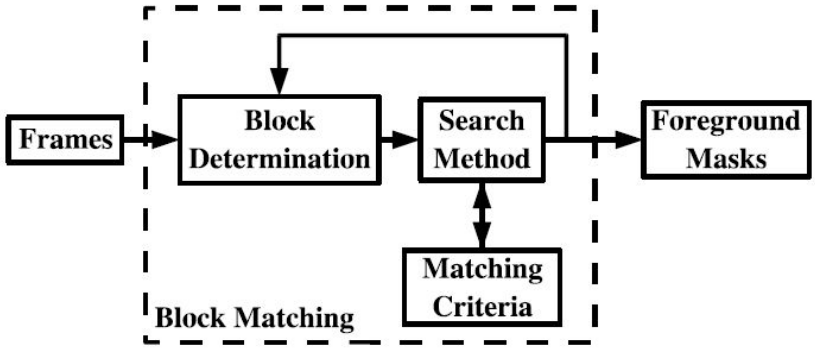


Fig 2. Block based matching approach [1]

In this project, we take a block of each pixel from reference image and scanned row-wise to match the other blocks by calculating their sum of square difference (SSD). We considered only those blocks where the SSD is minimum in their row. This algorithm is tested on different block size such as 3x3, 5x5 and 9x9 and we inferred that as the block size gets larger, the image becomes much more smooth.We set the search range as 50 by observing the given stereo images.

Dynamic Programming

Since window based approach is computationally expensive and leads to some problems like adaptive windowing etc., we solved this problem through Dynamic Programming. One of the applications is to find the longest common subsequence of a string (in this case image pixels). We take the optimal path while backtracking i.e. vertical and horizontal path for occlusion and diagonal path for objects.

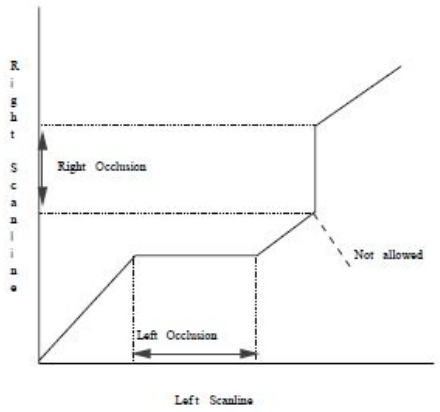


Fig3. Optimal path for backtracking[2]

The following constraints were considered while doing stereo imaging with dynamic programming

- Uniqueness constraint: From any point in one image there should be at most one point in the other image
- Monotonic constraint: The change in disparity is smooth unless encountered with depth discontinuities.

In the dynamic programming, we first create a cost matrix for every row of size NxN for the image of size MxN. For every iteration we checked the vertical, horizontal and diagonal elements and take the minimum of these. For the disparity, we considered only the diagonal elements as matching and left or right elements as occlusion. In this project we applied standard reference paper algorithm taking only basic uniqueness and monotonic constraints.

Software and other methodologies used in this project

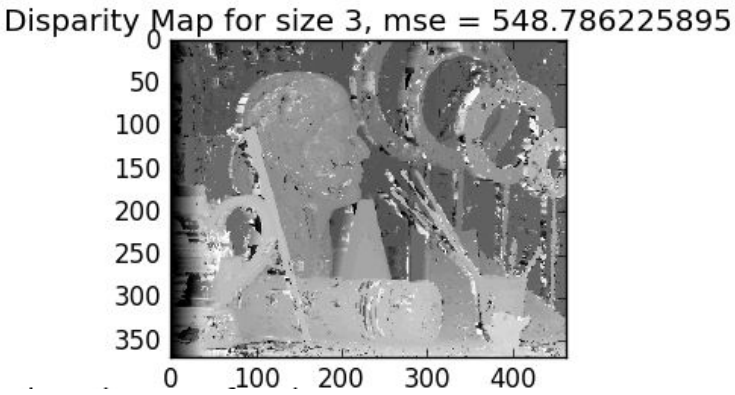
The project was implemented in python and we used Spyder IDE.

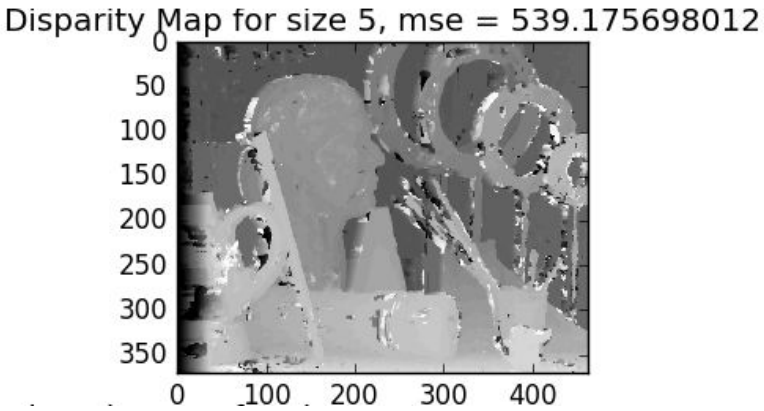
Outcome and Deviation

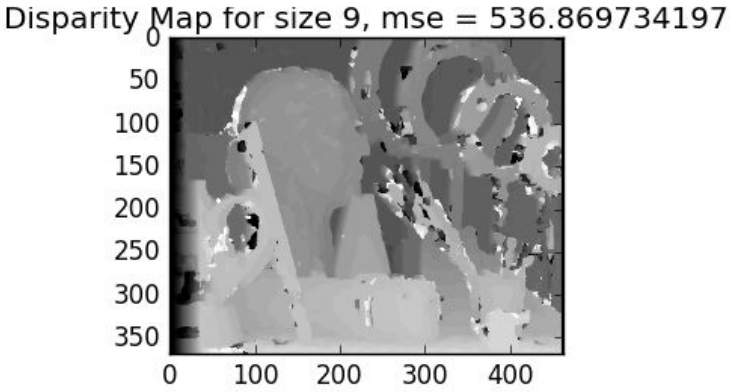
i) Presentation of Project Outcome

Block Based Matching

The following are the disparity based on different block sizes and their means square error. In general, we can see that the nearer objects are more brighter than the background. Since the pixels have higher value when the shift is larger, so brighter the pixel.

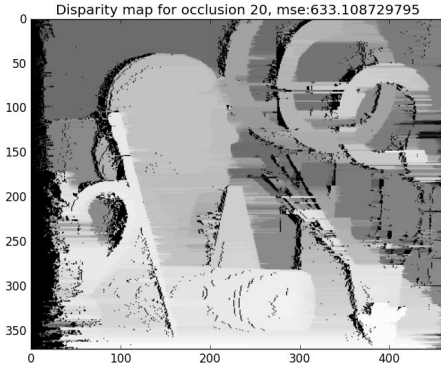


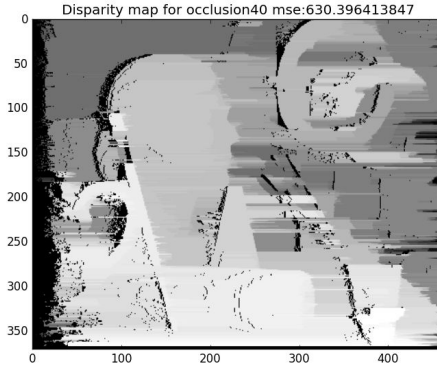


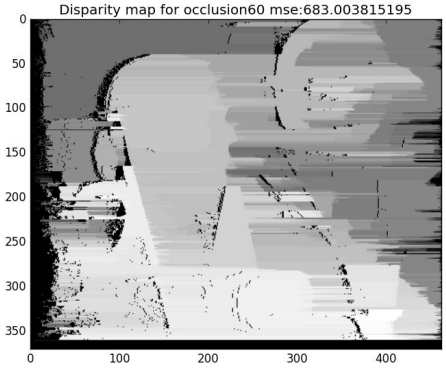


Dynamic Programming

The following are the disparity maps based on their occlusion value and their respective mean square error value. We tried different values of cost ranging from 20 to 60 but there were slight deviation in the results.







ii) Discussion of the outcome

1) We can infer from the block based matched results is that as the block size increase, the smoother image we get and value of MSE also decreases.
2) In Dynamic Programming there were slight deviation since there was slight improvement in mean square error value.

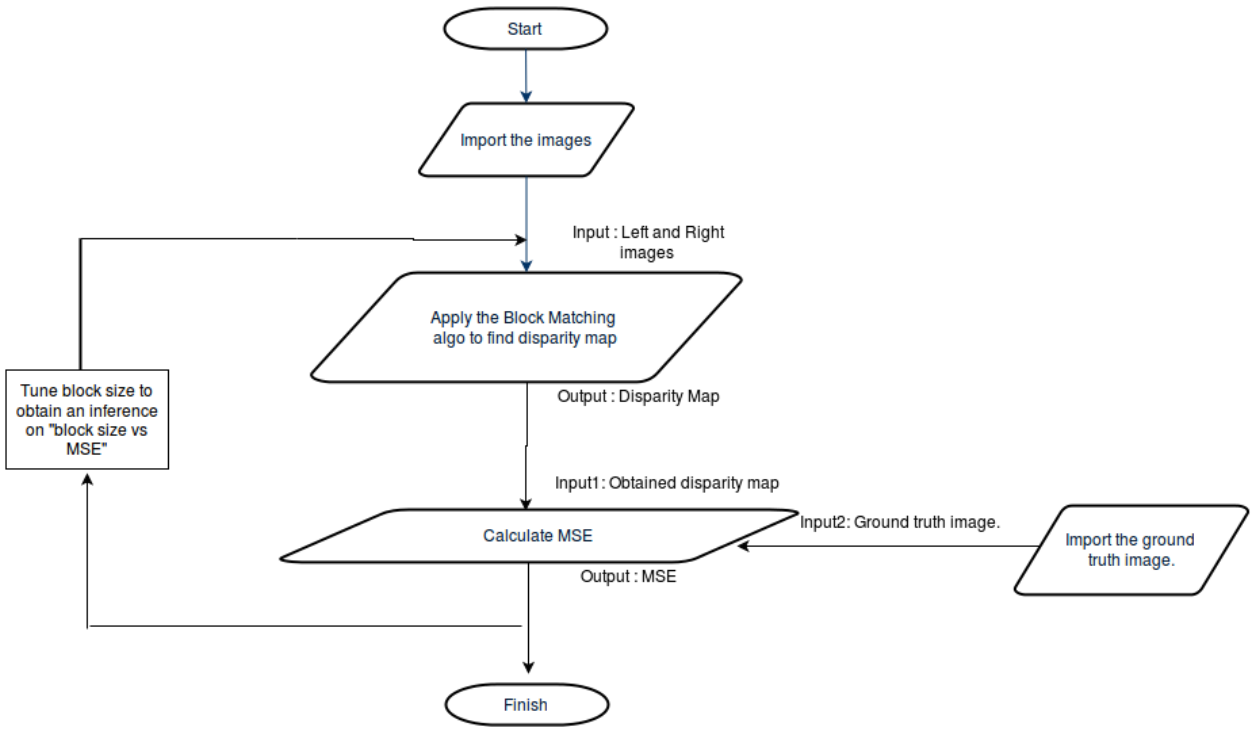
iii) Lessons learnt from algorithm

We learned how to calculate the disparity map by using block based matching and dynamic programming approach.

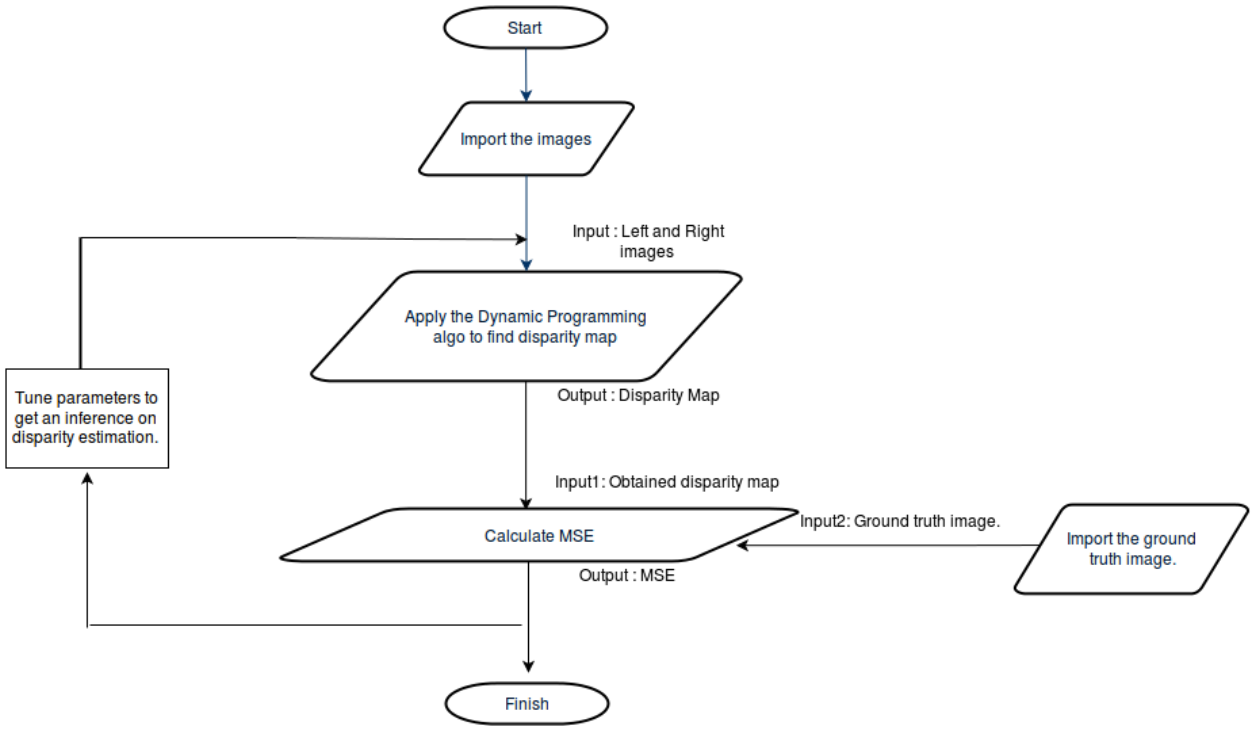
Explanation of Software and Program Development

We developed the software as per the below flowchart. We wrote the code in python.

Block matching:



Dynamic Programming



Summary and Discussion

i) Summary of the project

We calculated the disparity map from 2 rectified stereo images. Disparity map gave us the information of depth. The nearer objects were brighter in the results and the farther object were darker because of shift in the location of the pixels. Various parameters were tuned and such as block size in Block Based Matching Approach and Occlusion in Dynamic Programming. Block based approach gives more MSE than Dynamic Programming.

ii) Lessons learned in this course, including classroom,

We learnt various techniques pertaining to image processing and computer vision. These techniques, even though seemed simple prima facie turned out to be extremely powerful. Detecting objects in images was very interesting. We learnt how to detect edges, circles etc. Along with the theory, we learnt programming skills to solve basic image processing and computer vision problems

Acknowledgement

We are thankful for Prof Chen for having given guidance at every step o fthis course. The project could not have been completed without the TA's assistance. We thank the TAs as well.

References

[1] Nicole S. Love and Chandrika Kamath, An Empirical Study of Block Matching Techniques for the Detection of Moving Objects

[2] Ingemar J. Cox a, Sunita L. Hingorani a, , Satish B. Rao a, , Bruce M. Maggs, A Maximum Likelihood Stereo Algorithm,

[3] Lecturer Slides

[4] http://mccormickml.com/2014/01/10/stereo-vision-tutorial-part-i/