# State University of New York at Buffalo

## CSE 473/573 Summer 2016 Programming Assignment #2

Date: Tuesday June 28, 2016; Due: ***Wednesday Jul 13, 2016 at 3:00PM***

Problem (1) **(Stereo Vision)** 100%

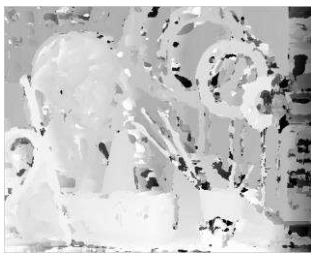
You are given with image dataset which has three views of the same scene and ground truth disparities of two views. In this project you will implement disparity estimation from rectified images using basic block matching and dynamic programming. You will also implement View Interpolation which is an interesting application of Stereo Vision.


Use the basic block matching techniques from class lectures to estimate disparity between view1 and view 5 of the image data. You will generate two disparity maps, one for view1 and the other for view2.

(i) Try different block sizes (3x3 and 9x9) and generate disparity maps. Report the estimated disparity map and the calculated MSE (Mean Squared Error) with respect to the given ground truth disparity maps. For images of **I** and **K** of dimension **m** x **n, MSE** is defined as

$$MSE = \frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

Below is a sample image of disparity map estimated using block matching. Eliminate the noisy disparity estimates using consistency checking – Due to occlusion, we have missing regions in the stereo image pair. This will result in random matches, since a correct match does not exist. After obtaining disparity maps for both the images, use back-projection to do a consistency check. Ideally, if the disparity maps are accurate, if you pick a pixel on left scan line and find its position on the right scan line, you would see the disparity value at that pixel is consistent i.e. that pixel should back-project to the pixel you chose on left scan line. Mark the inconsistent pixel with the value Zero --- 30%



(ii) Estimate disparity using Dynamic Programming by following the pseudocode from section 2.1 of http://raycast.org/powerup/publications/stereo.cvpr.pdf or http://www.cs.umd.edu/~djacobs/CMSC426/Slides/stereo_algo.pdf --- 40%

(iii) View3 is what you would see if you place a camera exactly at the midpoint of the baseline of cameras that captured view1 and view5. Your task is to generate view3 by using view1 and view5 images and their respective ground truth disparity maps. The synthesized view will have holes which can be marked as Zero. Below is the snapshot of the expected result --- 30%