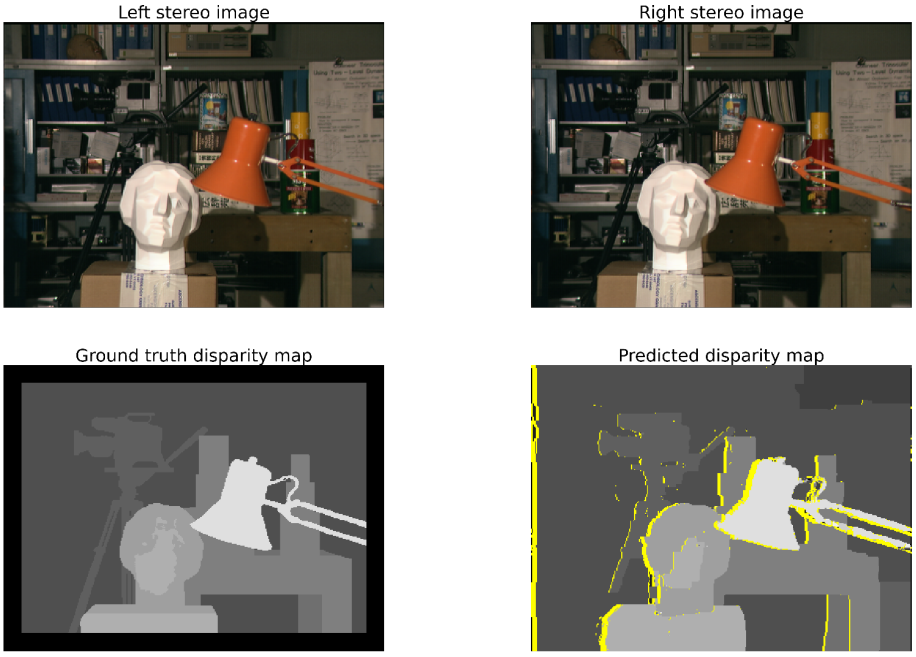
CV Assignment 4: Graph Cuts

Deadline: 14th March 2023, 23:59 hours

Concerned TAs: Triansh, Gurkirat

Disparity maps in stereo images

One of the early attempts to measure the relative depths of a scene by calculating disparity maps from a pair of stereo images using graph cuts was carried out by Vladimir Kolmogorov & Ramin Zabih in their paper, "Computing Visual Correspondence with Occlusions using Graph Cuts."
This method is based on finding a pair of pixels that potentially correspond (mentioned as assignments in the paper), providing a natural way to know occluded pixels between the images. You are supposed to read and implement the algorithm described in the paper.
You can refer to this publication to get in-depth detail about the implementation and concepts needed for the algorithm.
The final results of the algorithm will look similar to the below image. You can also refer to Figure 3 of the paper mentioned above.



Sample output for the given pair of stereo images. Yellow-colored pixels in the predicted disparity map denote the occluded pixels (visible in the left image but not in the right image). Y*our results may differ slightly based on your implementation.*

Evaluation Metrics

Once you have implemented the algorithm, you need to compute the following metrics:

1. Total time taken.
2. The percentage of occluded pixels in the disparity map.
3. *Total error:* The percentage of incorrectly identified disparities (Don't count occluded pixels)
4. *Gross error:* The percentage of pixels with incorrectly identified disparity but a disparity within ± 1 range of the correct disparity. (Don't count occluded pixels)

For metrics 3 & 4 refer to the table in section 6 of the paper.
Choose a total of 16 disparities when calculating these metrics.

Experiments

For each experiment, describe the results you observe briefly.

1. Include a study of how changes in the various parameters (in the below list) affect the results.
    a. Varying λ.
    b. Varying the maximum number of iterations.
    c. Varying the disparity range.
2. Compare your results by changing the data penalty function.
    a. The mean squared value of the difference in pixel intensities, i.e.,

$$d(x_L, x_R, I_L, I_R) = \sum_{c=1}^{3} |I_L(x_L)_c - I_R(x_R)_c|^2$$

    b. The Birchfield Tomasi Pixel Dissimilarity Measure. You can find the details for the formula in this paper by Stan Birchfield and Carlo Tomasi.

Implementation Details

You can use OpenCV functions for reading, writing, and basic manipulation of images. You are allowed to use libraries for the calculation of the maxflow/min-cut of the constructed graph. Using the maxflow package to calculate the maxflow of the constructed graph is **highly recommended**. Other networking/graph packages might give inconsistent results for the same.

The naive implementation of the algorithm can take a significant amount of time to run a single iteration. It is suggested to **avoid python for loops and list comprehensions** as much as possible. Alternatively, try to use numpy methods to speed up computations and reduce your waiting time. Some helpful numpy functions are already listed in the repository README.

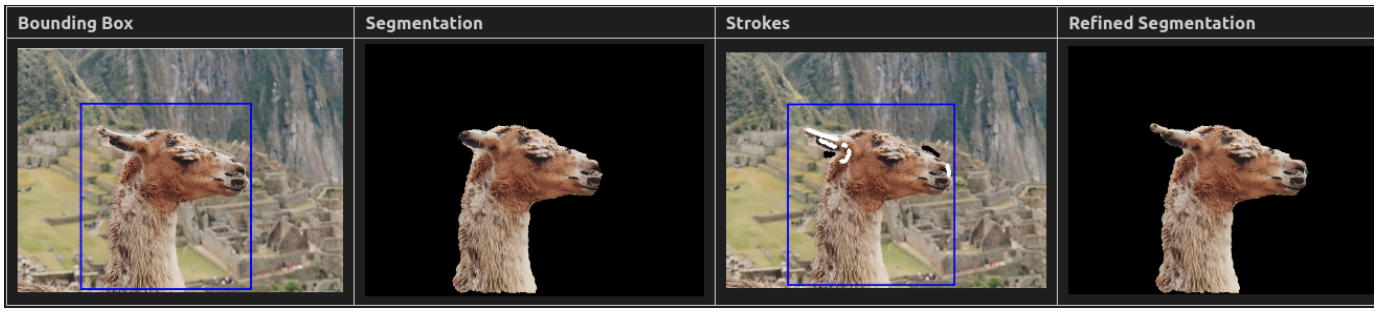
Please note that you are free to choose between the style of implementation, and you will not be graded based on the speed of your algorithm.

Please refer to the repository README for further instructions.

GrabCut

GrabCut algorithm was designed by Carsten Rother, Vladimir Kolmogorov & Andrew Blake from Microsoft Research Cambridge, UK. In their paper, "GrabCut: interactive foreground extraction using iterated graph cuts," they have developed an algorithm for single object foreground extraction with minimal user interaction.

You can use **OpenCV's implementation of Grabcut** for this part of the assignment.



Grabcut Algorithm, in a nutshell.

Experiments

In this task, you have to do the following experiments (on 3 images of your choice from the dataset) with GrabCut and evaluate the results using the evaluation metrics explained earlier

1. The number of iterations of GMM updating and energy minimization.
2. Effect of a tight initial bounding box or a loose bounding box.
3. Different color spaces or ways of representing pixels.

For each of the experiments, you need to give a brief description explaining the results you observe.

In addition to the above experiments, you need to report evaluation metrics values on at least 4 individual images: 2 images that the GrabCut algorithm performs very well on and 2 images that algorithm does not perform very well (if there are any).

Evaluation Metrics

1. *Accuracy*: The number of pixels that are correctly labeled divided by the total number of pixels.
2. *Jaccard similarity*: The intersection over the union of your predicted foreground region with the ground truth.
3. *Dice similarity coefficient*: It is 2 * the Area of Overlap divided by the total number of pixels in both images.

You can refer to this article to understand these metrics.