

MSML640 Assignment 4

Ayush Chaudhary (120429125)

April 10, 2024

1 Short Answer Problems

1. When using the Laplacian of Gaussian (LoG) for interest point detection:
 - (a) Because the Laplacian of Gaussian operator is a scale-space operator, the results vary when selecting the spots with the highest response at various scales. This implies it reacts differently to characteristics on various scales. Points recognized at smaller scales often capture finer details and minor characteristics, whereas points detected at bigger scales capture wider aspects. As a result, picking sites with the highest reaction at different scales enables the detection of objects with varying levels of detail and size.
 - (b) If we just chose points with a strong reaction without considering the scale, we risk missing essential traits that appear at different scales. This can lead to information loss and, consequently, worse detection quality. By taking into account the response scale, we can ensure that characteristics of various sizes are appropriately recorded. This increases the overall quality of identified points by presenting a more complete representation of the image's characteristics at various sizes.
2. In RANSAC for stereo vision with uncalibrated views, an inlier is a correspondence between points in two views that is near to the epipolar line predicted by the calculated fundamental matrix. To calculate inliers, we randomly choose pairs of correspondences, estimate the basic matrix, compute epipolar lines, and count correspondences that are within a certain distance of their expected lines. We continue this process iteratively, choosing the basic matrix with the greatest inliers.
3. Two possible failure modes for dense stereo matching using local appearance and correlation search within a window are:
 - (a) Occlusions: When an object in one view is obscured by another object in the other view, the correlation search may fail to find a match. This is because the obscured object's appearance is not visible in the other view, leading to a lack of correlation between the two views.
 - (b) Textureless regions: In regions with little or no texture, such as a blank wall or a uniform surface, the local appearance and correlation search may not be able to find a reliable match. This is because the lack of distinctive features or patterns makes it difficult to establish correspondences between the two views.
4. The value recorded in a single dimension of a SIFT keypoint descriptor represents the gradient magnitude and orientation of the image at that point. The descriptor is divided into 16 sub-regions, each of which is further divided into 4 sub-regions. The gradient magnitude and orientation are calculated for each sub-region, resulting in a total of 128 values in the descriptor. These values are used to describe the local appearance of the image around the keypoint and are used for matching and recognition tasks.
5. While SIFT extracts image features and Hough Transform detects specific shapes in its own parameter space, they complement object recognition by providing complementary information: SIFT for feature matching and Hough Transform for potential geometric cues from the object. They don't directly define a shared parameter space.

2 Programming

1. Results for Crop images



Figure 1: Actual Crop Images

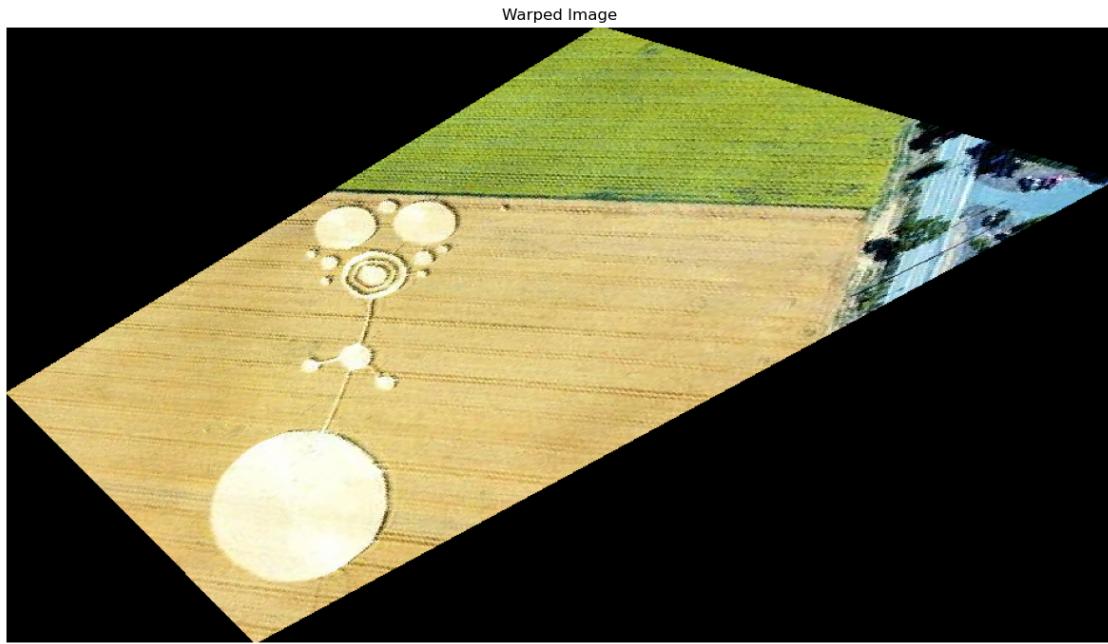


Figure 2: Warped Crop Image

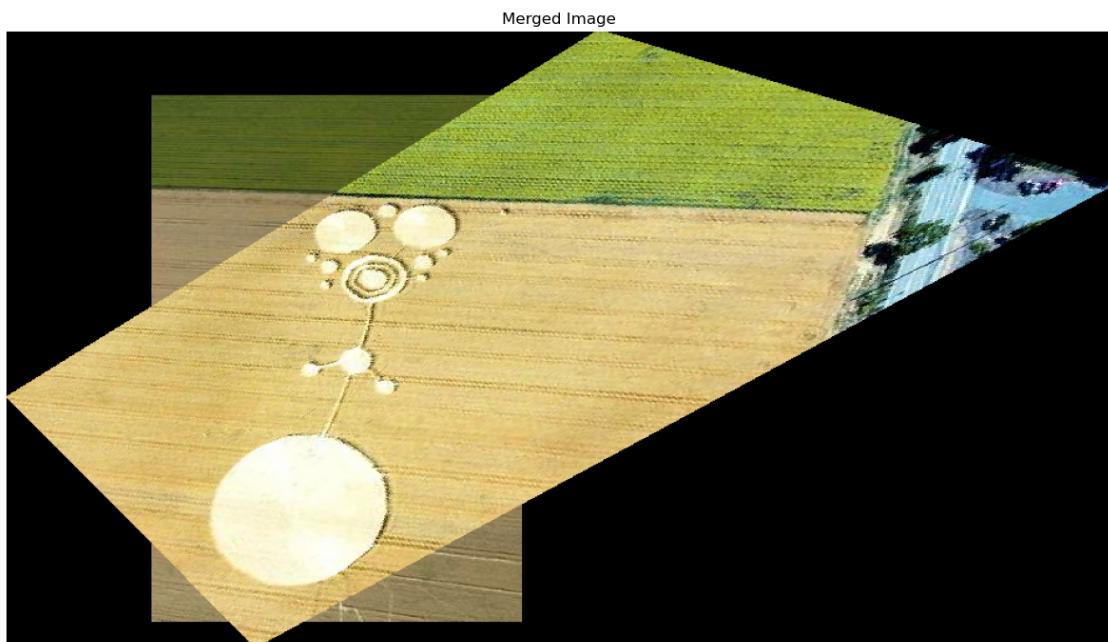


Figure 3: Merged Crop Image

2. Results for WDC Images

Image 2

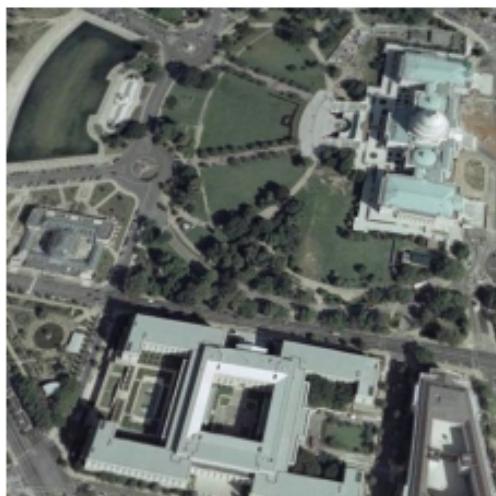


Image 1



Figure 4: Actual WDC Images

Warped Image

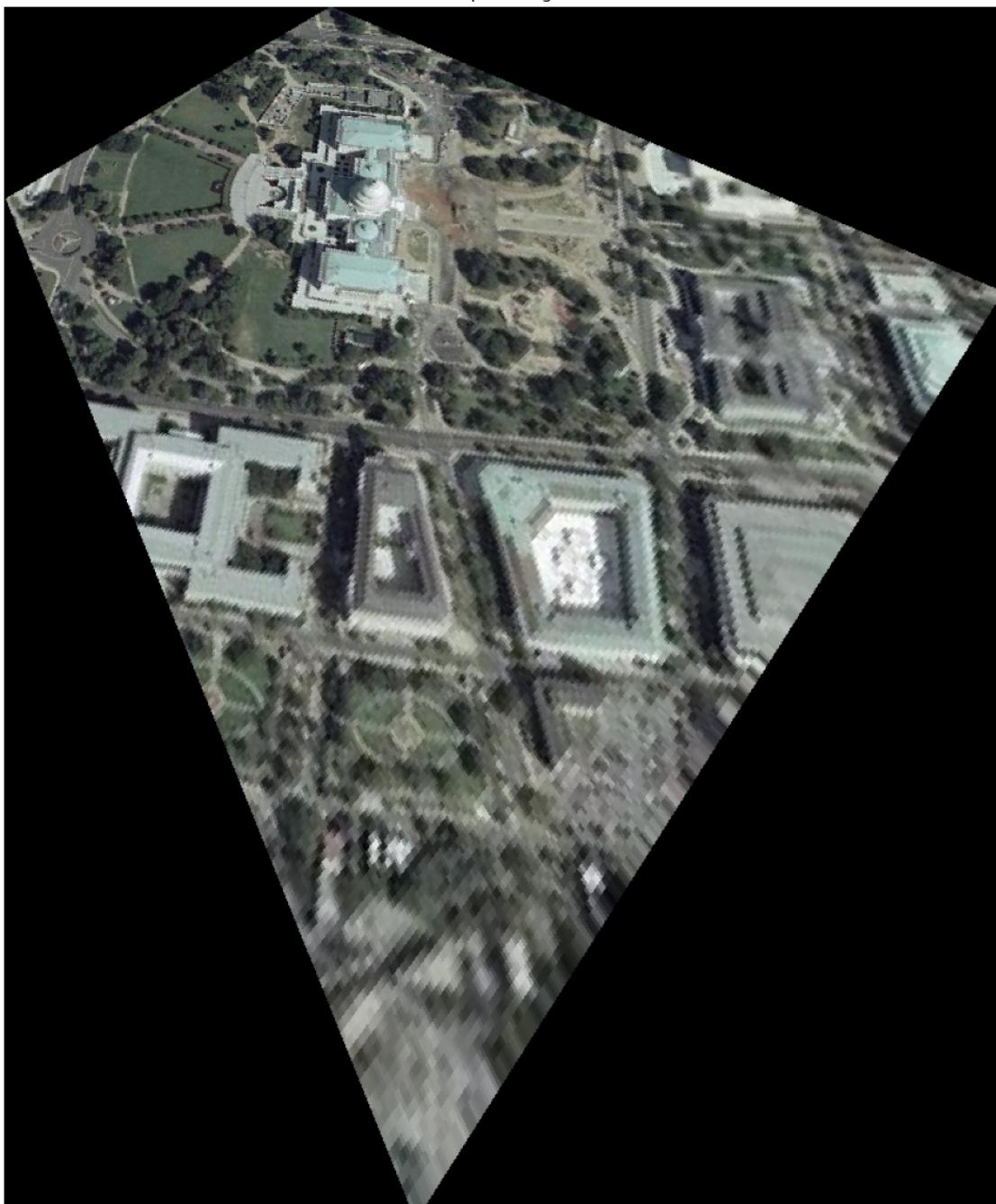


Figure 5: Warped WDC Image

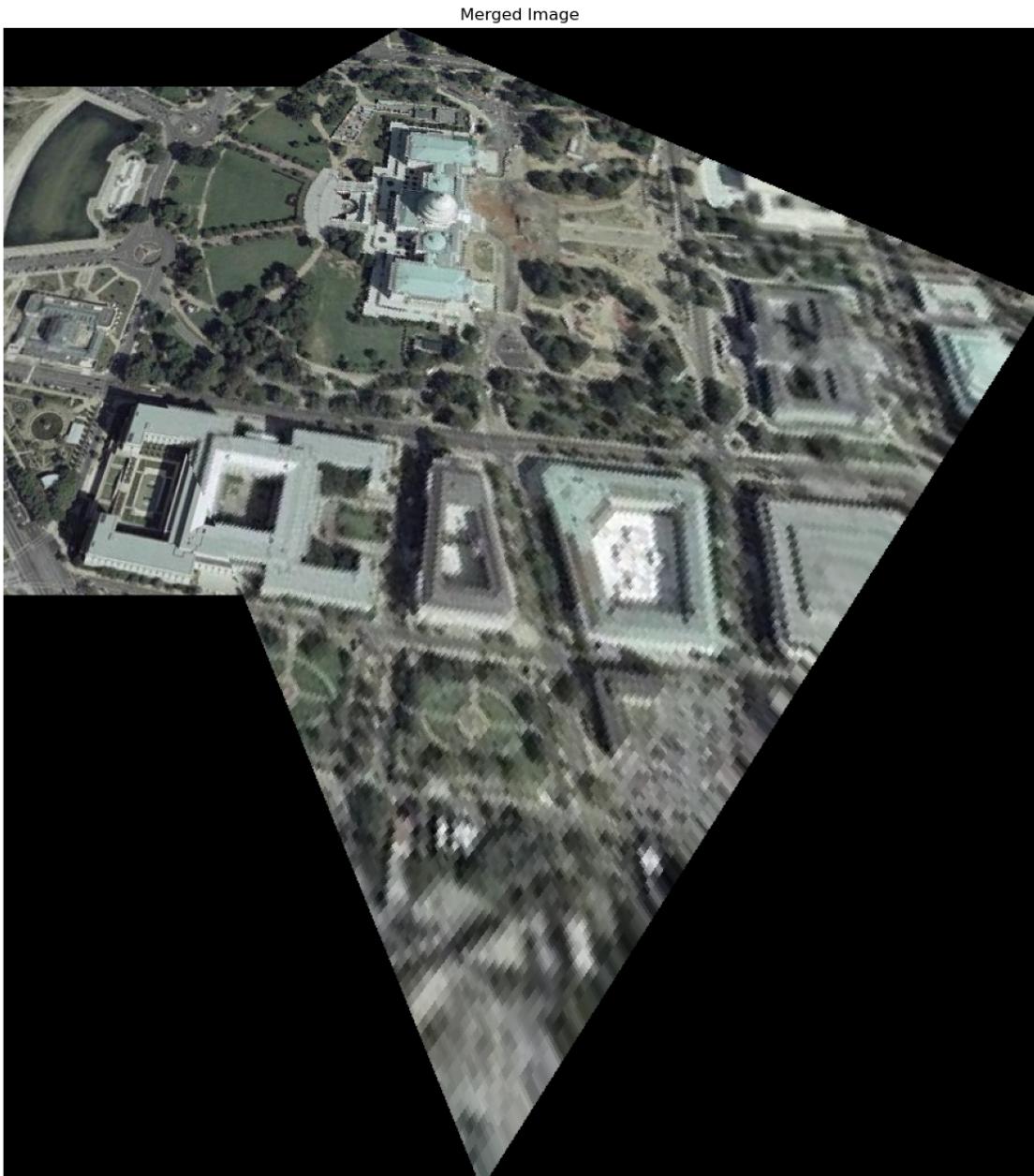


Figure 6: Merged WDC Image

3. Results for custom images

Image 1



Image 2



Figure 7: Actual Custom Images

Warped Image



Figure 8: Warped Custom Image

Merged Image



Figure 9: Merged Custom Image