Maastricht University

Department of Advanced Computer Science

Data Science for Decision Making

KEN4225 - Computer Vision
Resit First Assignment

Author:
Christos Koromilas

Student Number:
i6345703

July 8, 2023

1 Introduction

The application of image feature matching and transformation is critical in several computer vision tasks, including image stitching, object recognition, and 3D reconstruction. This assignment delves into the detailed process of feature matching and transformation estimation between two images utilizing the Harris corner detection algorithm, SIFT descriptors, and the RANSAC algorithm.

2 Image Acquisition and Pre-processing

This stage of the image stitching procedure involves the acquisition and pre-processing of the images. The 'left' and 'right' images are read into the program with OpenCV's imread() function, which allows the images to be imported as pixel arrays, preserving their original RGB color information.

Upon acquisition, the color images are converted to grayscale, a common pre-processing step in many computer vision tasks. This conversion simplifies the images, decreasing computational complexity while retaining critical visual information, thereby facilitating the subsequent feature detection and matching stages.

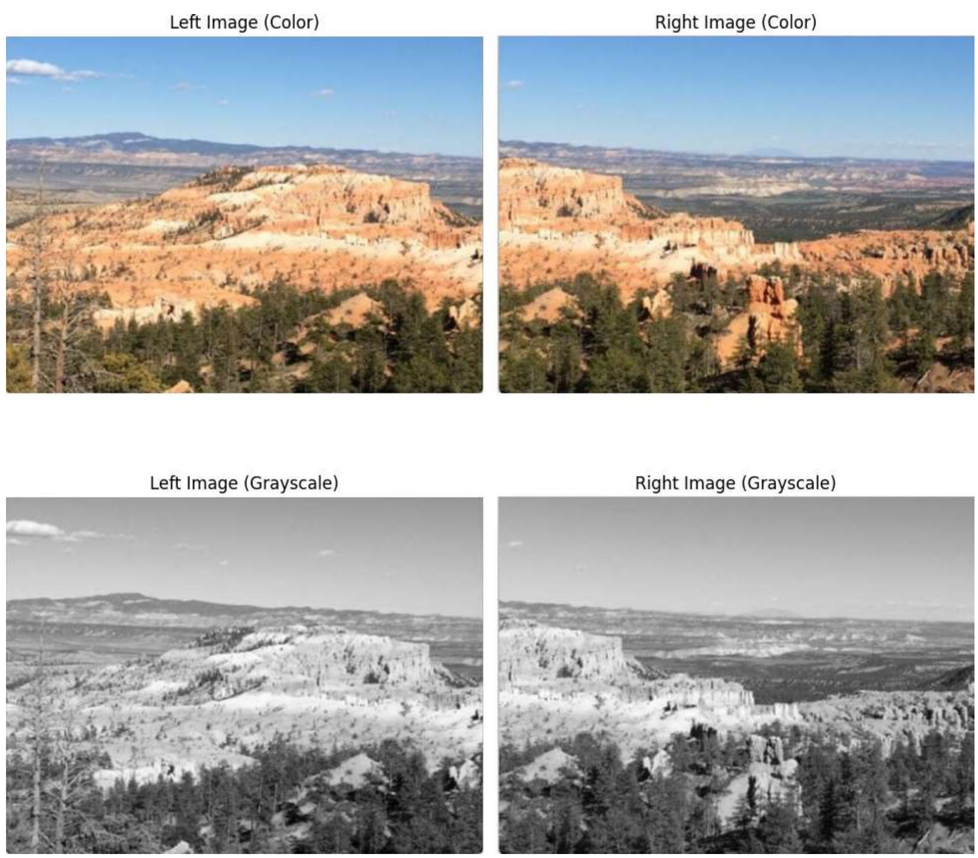


Figure 1: Pre-processing: From RGB to Grayscale

3 Feature Extraction: Harris Corners

The succeeding step of the process is the extraction of features from the grayscale images. This stage is crucial as it enables the identification of points of interest, which are pivotal for precise image alignment.

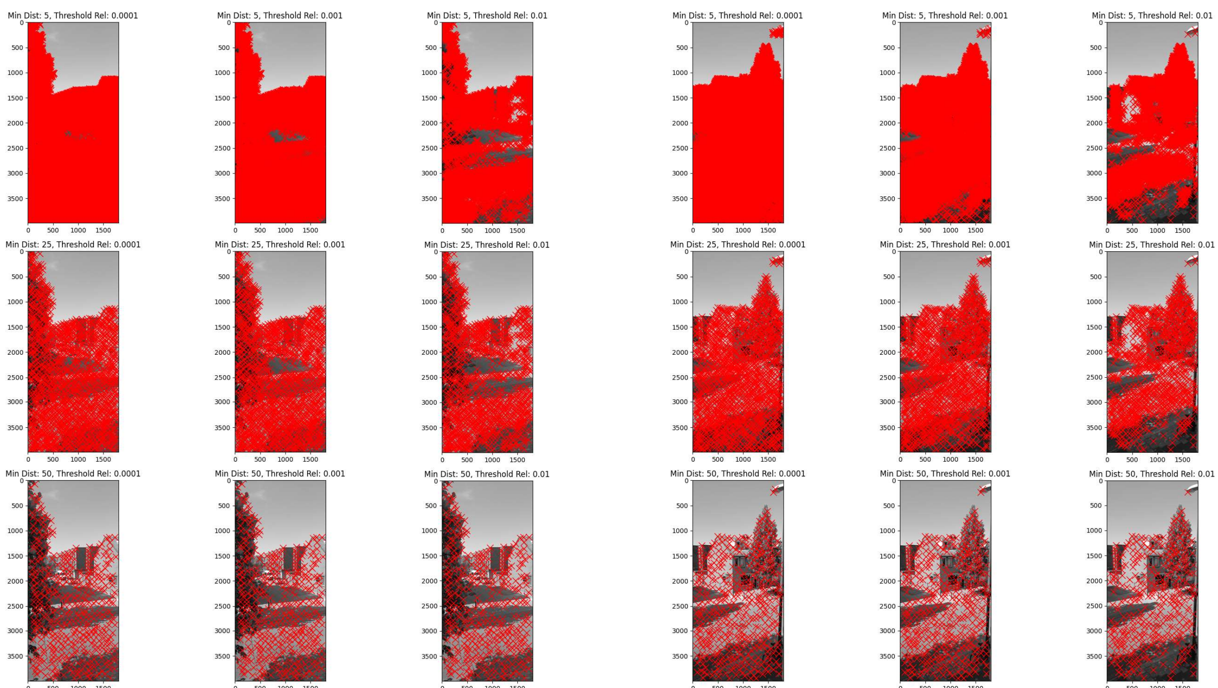
3.1 Corner Detection

The Harris Corner Detection algorithm, a prevalent technique for identifying distinctive corner points in an image, is employed for feature detection. The corner-harris() function calculates the Harris corner measure response for each pixel in the grayscale image, identifying the corners.

To optimize the matching process, the corner-peaks() function finds the peak of the corner measure response map, selecting corner points with a minimum distance and relative threshold. Additionally, the orientations of the detected corners are computed using the corner-orientations() function, providing extra information for the matching process.

Sensitivity Analysis Visualizing the detected Harris corners enhances the understanding of the corner detection process. A sensitivity analysis is conducted to determine the optimal parameters for the Harris corner detection. Various combinations of the minimum distance between corner points and the relative threshold are tested, visualized, and the combination yielding the most accurate and distinct corner points is chosen.

Upon the selection of the optimal parameters, the Harris corner detection is rerun, providing a more refined set of corners for each image. The final step of this stage is to visualize the detected corners and provide the count of corners detected in each image for debugging and further optimization of the algorithm.



(a) Harris Corners Detection: House Left

(b) Harris Corners Detection: House Right

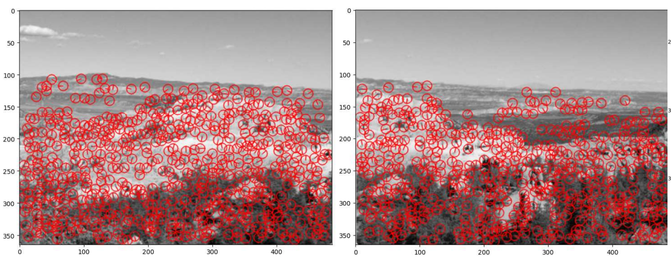
Figure 2: Sensitivity Analysis Results

3.2 Extracting Fixed-Size Patches Around Corners

This step involves the extraction of fixed-size patches around the corners identified in the previous step. These patches, along with their corresponding orientations, are represented as OpenCV KeyPoint objects for compatibility with later processing steps. From the sensitivity analysis above and for different patch sizes, the perfect matches, taking too small or too big sized patches around the corners and can lead to errors during the final transformation. After experiments on different kinds of images, a patch size = 15 gets the job done relatively good.

3.2.1 Scale-Invariant Feature Transform (SIFT)

SIFT is utilized to compute descriptors for the extracted patches, capturing local image features around each keypoint. This information facilitates the comparison of corresponding regions between different images. The generated SIFT descriptors are then visualized on the original images.



(a) SIFT Descriptors: Classic

(b) SIFT Descriptors: House

Figure 3: SIFT Descriptors Visualized

4 Calculation of Euclidean Distances

Subsequent to the computation of the SIFT descriptor, the Euclidean distance between every couple of normalized descriptors, derived from the left and right images, is calculated. These Euclidean distances function as indicators of the degree of similarity between pairs of descriptors, with smaller distances signifying highly similar descriptors. This is the heatmap of the distances for 2 images (left and right).

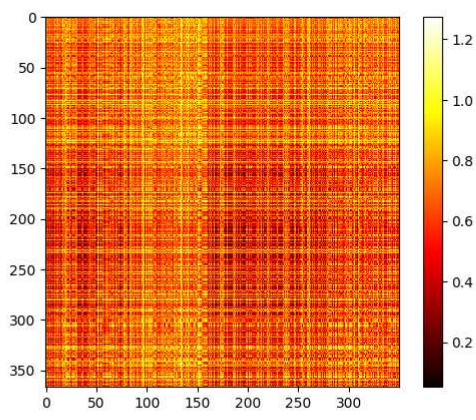


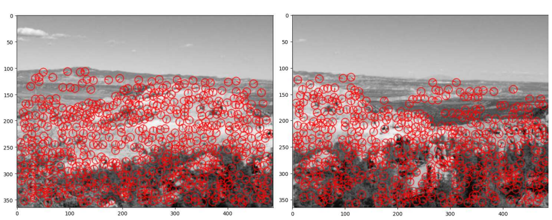
Figure 4: House: heatmap

5 Feature Matching Process

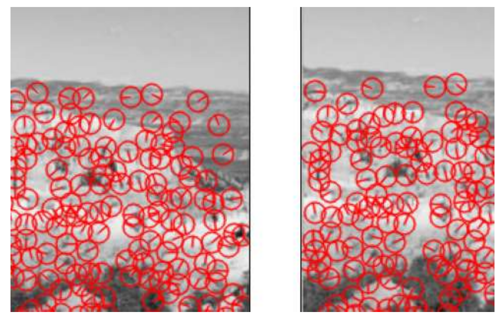
5.1 Selection of Matches

Following this, the computed Euclidean distances are sorted and the pairs of indices corresponding to the smallest distances are selected. These pairs represent the pairs of descriptors from the two images that exhibit the highest similarity. To maintain the uniqueness of correspondences, a pair is only chosen if its keypoints have not been selected previously.

Post the selection of the most apt matches, they are visualized on the original images. This step not only confirms the accuracy of the match selection process but also provides a visual representation of the correlations between the two images.



(a) Classic with the descriptors



(b) Classic zoomed

(c) House with the descriptors

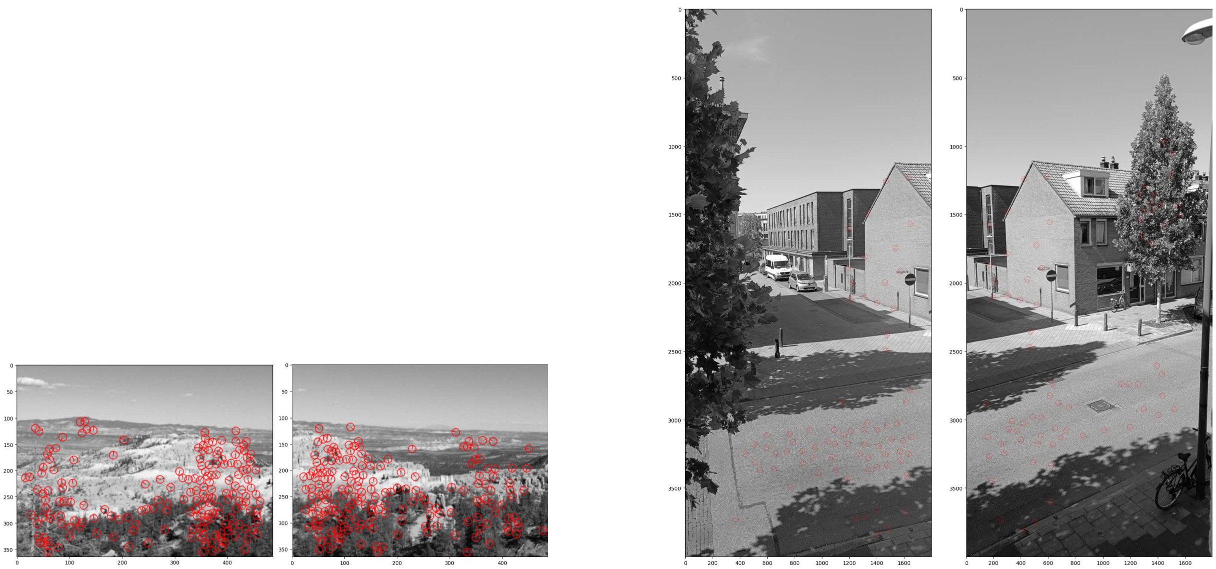
(d) House zoomed

Figure 5: Descriptors and zoomed versions of the images

It is understandable that the majority of pairs that aren't selected can be located on parts of the images that don't overlap. This typically refers to the left part of the left image and the right part of the right image. For instance, we observe that most points not selected appear on the right side of the right image, which isn't featured in the left image.

6 Utilization of RANSAC

The selected matches are employed to determine the best affine transformation between the two images by utilizing the Random Sample Consensus (RANSAC) algorithm. This algorithm iteratively estimates the transformation using a randomly selected subset of matches and evaluates the quality of the estimation based on its alignment with all matches.



(a) Classic

(b) House

Figure 6: RANSAC

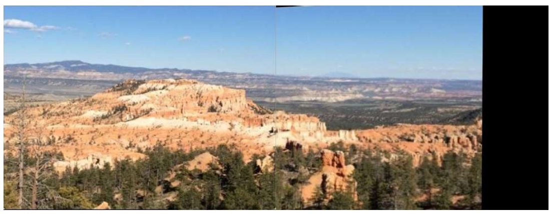
7 Estimation of Homography

In the final stage, the homography between the two sets of keypoints is computed utilizing the matches derived from the RANSAC algorithm. The homography matrix encapsulates the optimal transformation required to align the two images for stitching. This facilitates the seamless merging of the two images to generate a panorama based on these matches.

8 Warping and Stitching of Images

In this segment of the process, the focus is on image warping and stitching. Image warping is a technique employed to adjust images to account for disparities in camera perspective. Leveraging the previously computed homography matrix, H, a perspective warp is performed on the right image to align it accurately with the left image.

Following the warping process, the left and right images are stitched together. This is accomplished by generating a new image capable of accommodating both the warped right image and the original left image, followed by the integration of the left image into this newly constructed image canvas.



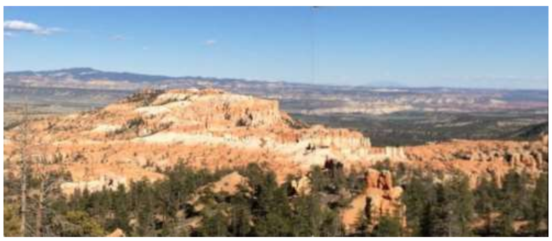
(a) Classic

(b) House

Figure 7: Wraped images

9 Image Cropping and Presentation

Upon successful stitching of the images, the final stage of the process focuses on image cropping to enhance the aesthetics of the image. The stitched image might have black regions due to the warping and stitching processes. To ensure a cleaner and more focused result, I locate the bounding box that encapsulates all non-zero pixels in the stitched image and crop the image to these dimensions.



(a) Classic

(b) House

Figure 8: Final Images

10 Assessing Accuracy

Subsequently, we require a metric to evaluate the efficacy of our results. For this purpose, we calculate the Euclidean distance between the chosen key-point coordinates in Image 1 and the corresponding transformed ones in Image 2. A sensitivity analysis is then performed to understand how different parameters affect this score. However, it should be noted that this score is only relative for the same fixed-size patches experiments and cannot be compared across experiments with different sized patches. For the image with the house: 325 and for the classic: 47.

11 Conclusion

The conducted study effectively demonstrated the application of algorithms for feature matching and image stitching, which are key components of computer vision. Through various stages including image acquisition, pre-processing, feature extraction, feature matching, and image transformation, the assignment successfully stitched together two different images, thereby generating a panorama.

Key points of interest in the images were identified using the Harris Corner Detection algorithm, and their corresponding descriptors were computed using the Scale-Invariant Feature Transform (SIFT). Matches were then selected based on the Euclidean distances between descriptor pairs. The RANSAC algorithm was utilized for selecting the most apt matches and estimating the best affine transformation.

The computation of the homography facilitated the alignment of the images. The images were then warped and stitched together using this homography. A final cropping step helped to enhance the visual appeal of the final stitched image.

It's clear that image stitching has vast potential applications in a variety of fields, including surveillance, navigation, virtual reality, and panoramic photography. However, achieving successful results involves careful parameter tuning and algorithm selection, as revealed by the sensitivity analysis of the Harris corner detection and patch size determination.