RBE 549- Project 1:MyAutoPano

Ankit Mittal
Department of Robotics Engineering
Worcester Polytechnic Institute
Email: amittal@wpi.edu

Rutwik Kulkarni
Department of Robotics Engineering
Worcester Polytechnic Institute
Email: rkulkarni1@wpi.edu

***Abstract*—In this project, we propose techniques for stitching images to generate seamless panoramic scenes. The process involves using pairs of images with overlapping regions. Our objective is to merge these images into a single panoramic image. Our approach encompasses both traditional computer vision methods and deep learning techniques. The classical method we've used is a feature-based technique that establishes a relationship between images by identifying and matching features. For the deep learning component, we employed both supervised and unsupervised learning strategies to estimate the homography, which is the transformation between a pair of images, to achieve the desired output.**

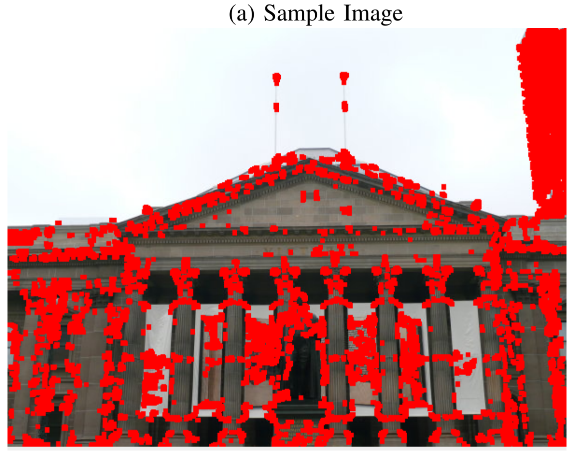
I. Phase 1: Classical feature based technique

Feature-based methods focus on identifying features present in the overlapping areas of images. These features, known as keypoints, are matched between the images. We use this matching to estimate the homography, which represents the transformation between these sets of points from the two images. This homography is then applied to warp the second image to align with the first. Following this alignment, the images can be seamlessly stitched together, using the first image as the reference. Each subsection of our documentation provides detailed explanations of the methodologies used and the resulting outputs of the images.

A. Corner Detection

The concept centers on establishing connections between images by identifying a set of features. Corners are particularly effective for this purpose as they are discernible from various viewpoints. By detecting as many corners as possible in a given image, we can compare these features between images. This comparison provides insights into the geometric relationship between the images. For corner detection, we utilized the Harris corner detection function available in OpenCV. We determined the following parameters to be optimal for corner detection: a kernel size of 3, a Harris K parameter of 0.04 Figure 1 illustrates the corners detected in the images using these parameters.

(a) Sample Image



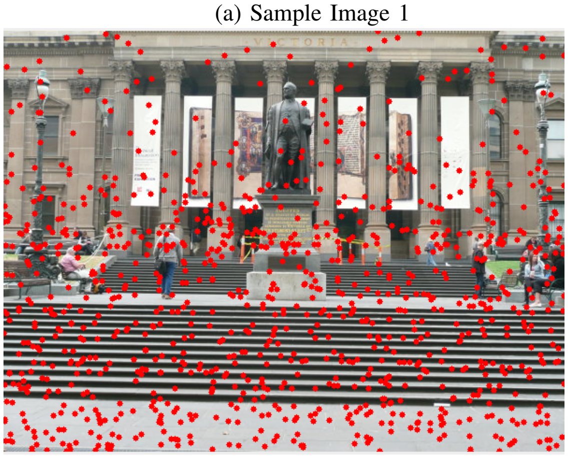
(b) Harris Corners

Fig. 1: Image corner detection

B. Adaptive Non Maximal Suppression (ANMS)

Upon detecting corners in each image, the challenge is to identify the "best" corners, characterized by their distinctiveness among local peers and even distribution across the image for superior homography. This is addressed using Adaptive Non-Maximal Suppression (ANMS), which operates by first selecting local maxima among the corners, focusing on those farthest from stronger corners. The essence of ANMS is its effectiveness in evenly distributing corners, countering the tendency of corner detectors like the Harris method to identify clusters of corners, rather than individual ones. This is due to corners encompassing multiple pixels, particularly in high-resolution images. ANMS overcomes this by choosing points maximally distant from stronger corners, ensuring that when the top N best corners are selected, they represent individual points from each cluster. Figure 2 illustrates the transformation from clustered to well-distributed, optimal corners achieved through ANMS.

(a) Sample Image 1



(b) Sample Image 2

Fig. 2: Image corner After ANMS

C. Feature Descriptor

To effectively compare corners across images, we assign them unique identities using Feature Descriptors. In line with our approach outlined in the problem statement, we first select corners that fit well within a 40x40 dimension. These selected corners are then flattened into a 1D array. To create a representative sample, we pick pixels at every 25th index, essentially capturing every 5th pixel in a row-wise or column-wise manner. This process results in a patch of dimensions 8x8. We standardize and blur this patch to ensure a smooth variation, providing each corner with a distinct and comparable identity, encapsulated in its Feature Descriptor. This methodology allows for precise comparison and alignment of corners across different images.

D. Feature Matching

The feature descriptors obtained through our process are then matched with descriptors from other images, a crucial step as it indicates the extent of overlap between the images. Ideally, photometric comparisons would provide the most accurate measure of overlap, but these methods are computationally expensive. Instead, we rely on geometric features like corners. Our comparison algorithm is based on David Lowe's ratio test.A simplistic approach would have been to minimize the distance between feature descriptor vectors, but this risks false positives leading to inaccurate matches. The ratio test circumvents this by comparing the distances of the first and second best feature descriptors. If the ratio of these distances falls below a certain threshold, the pair is considered a good match and is added to the feature match set. This process ensures the uniqueness and significance of each match. If the second-best match is comparably close to the first, indicating a lack of uniqueness, the pair is discarded. This method effectively screens for false positives, providing a straightforward and efficient way to match features across different data sets.

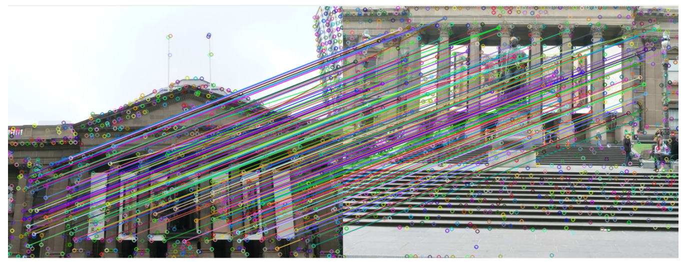


Fig. 3: Feature matching

E. RANSAC

The feature matches we've identified, like any data set, are susceptible to inaccuracies and outliers. To distinguish between valid data and outliers, it's necessary to employ a data model to fit and analyze the data. While manual pruning or using standard deviation and quartiles for data selection are possible methods, they require extensive tuning and may not generalize well across different data sets. Instead, we utilize Random Sample Consensus (RANSAC), a robust method that provides a probabilistic approach to creating a dataset free of outliers. RANSAC helps in identifying the minimal subset of data that best fits our model and generalizes effectively to the rest of the data points. In our context, the model in question is the homography estimation between a pair of images. Applying the RANSAC algorithm, we are able to refine the feature matches.

F. Blending Images(Warping and Stitching)

In our project, we developed a process to stitch two images together using a homography matrix. This process begins by assessing the dimensions of each image and identifying their respective corner points. Subsequently, we apply a perspective transformation to one of the images using the derived homography matrix. The next step involves creating a composite canvas, sized to accommodate both the original and the transformed images. This step is crucial as it involves translating the images to ensure all coordinates are positive, thus preventing any part of the images from being lost. Once the canvas is prepared, the function warps the second image in accordance with the pre-calculated transformation and integrates it onto the canvas. The final and critical step in this process is the overlaying of the first image onto the warped version of the second image. This results in a seamless and coherent panoramic image. Our method is particularly effective in merging images that have overlapping regions, thereby producing a single, expansive image. This approach has proven to be an efficient solution for creating panoramic views from multiple images.

Fig. 4: Image 1 & 2 Blending

G. Results

In our test set, we encountered some images that either contained irrelevant content ('garbage') or had very minimal overlap with other images, resulting in few or no feature matches. To address this issue, we implemented a threshold criterion for feature matching. We established that if the number of feature matches between a pair of images fell below 10, we classified the image as unsuitable for stitching. Consequently, such images were excluded from the stitching process. This approach ensured that only images with sufficient overlap and relevant content were considered, thereby enhancing the overall quality and coherence of the stitched panorama. Then we also tried the graph approach where tried to find the best matching image(having maximum overlap). Computing this on nomal image size is computationally expensive and for the best matching image for stiching we rsized all the images in the lower resolution. To further refine our image stitching process, we explored a graph-based approach aimed at identifying the best matching image, defined as the one with the maximum overlap. Recognizing that computing matches on full-sized images is computationally intensive, we adopted a strategy of resizing all images to a lower resolution. This resizing step significantly reduced the computational load, enabling a more efficient determination of the most suitable image for stitching. By focusing on lower-resolution images, we were able to swiftly and effectively identify the image with the highest degree of overlap, which is critical for achieving a seamless and high-quality stitched panorama.

Fig. 5: Set 1 Blending

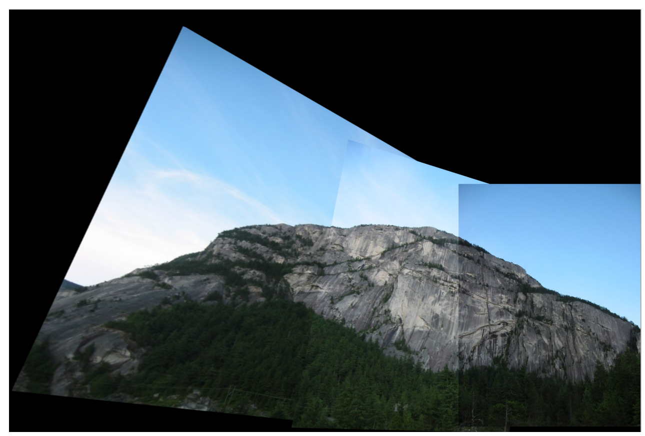


Fig. 6: Set 2 Blending

Fig. 7: Set 3 Blending

Fig. 8: TestSet 1 Blending

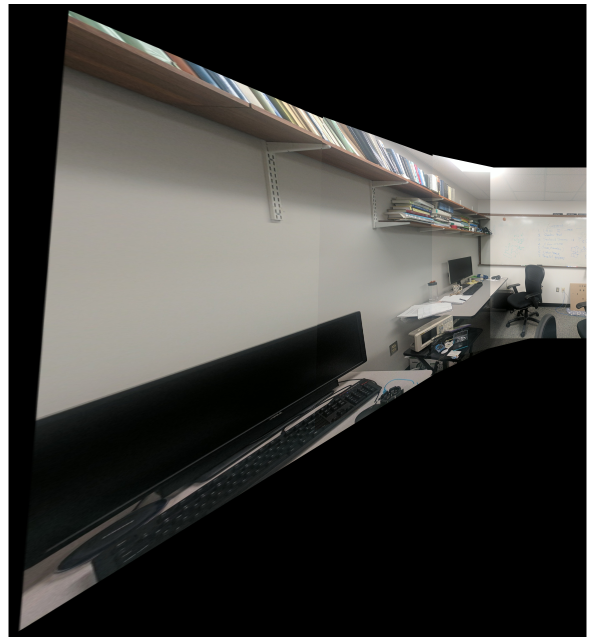


Fig. 9: TestSet 2 Blending

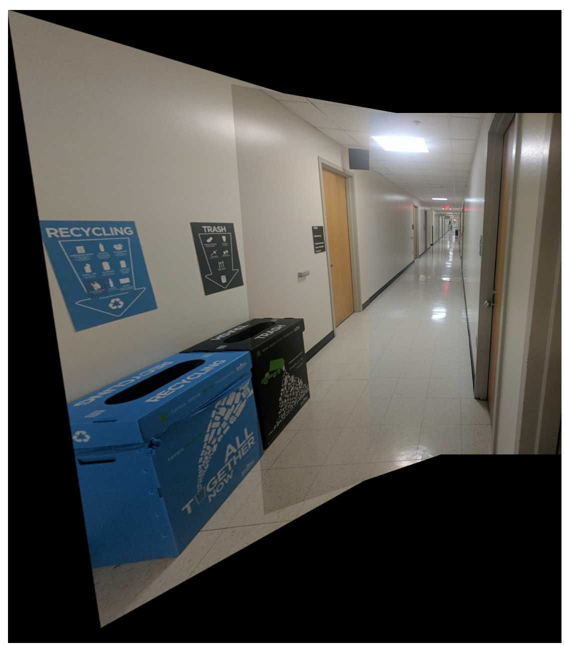


Fig. 10: TestSet 3 Blending