Automatic Panoramic Image Stitching using Invariant Features

Zenith (M25CSA032)
Department of Computer Science
CSL7360: Computer Vision

February 9, 2026

Abstract

This report presents an implementation of automatic panoramic image stitching, a fundamental problem in computer vision. We benchmark a robust approach based on SIFT features and Homography estimation against a naive concatenation method. Our experiments demonstrate that while naive stitching fails under camera rotation and perspective changes, the invariant feature-based approach successfully aligns images. We further show that a center-anchored stitching strategy minimizes cumulative distortion compared to sequential daisy-chaining.

1 Problem Statement

The objective is to combine multiple images with overlapping fields of view into a single high-resolution panorama. This requires:

- Detecting distinctive features invariant to scale and rotation.
- Establishing correspondences between overlapping images.
- Estimating the geometric transformation (Homography) to align them.
- Blending the aligned images to remove visible seams.

2 Methodology

2.1 Input Data

The dataset consists of three overlapping images captured with a handheld camera (Figure 1). The sequence covers a wide field of view with significant overlap between adjacent frames.

(a) Left View (b) Center View (c) Right View

Figure 1: Input images used for stitching. Note the overlapping fields of view and rotational perspective differences.

2.2 Feature Detection (SIFT)

We utilized the Scale-Invariant Feature Transform (SIFT) to extract keypoints. SIFT is chosen for its robustness to:

- **Scale**: Using Difference-of-Gaussians (DoG) feature space.
- **Rotation**: Assigning dominant orientations to keypoints.
- **Illumination**: Using gradient-based descriptors.

2.3 Feature Matching

Matching was performed using the FLANN (Fast Library for Approximate Nearest Neighbors) matcher with KD-Trees for efficiency. To Reject false positives, we applied **Lowe's Ratio Test**:

$$\frac{d(NN_1)}{d(NN_2)} < 0.7 \tag{1}$$

where $d(NN_1)$ and $d(NN_2)$ are the distances to the nearest and second-nearest observations. This filters out ambiguous matches, such as repetitive textures.

2.4 Robust Homography Estimation (RANSAC)

We model the transformation between two images as a planar Homography $H \in R^{3\times 3}$. To estimate H robustly in the presence of outliers, we employed **RANSAC** (Random Sample Consensus):

1. Randomly select 4 point correspondences.
2. Compute the exact homography H_{est} from these 4 points.
3. Count inliers: points where $||x' - H_{est}x|| < \epsilon$.
4. Repeat N times and choose the H with the most inliers.

2.5 Stitching Strategy: Center-Anchored

A distinct improvement in our implementation is the **Center-Anchored** strategy.

- **Naive (Daisy-Chain)**: *Left* $\to$ *Center* $\to$ *Right*. Errors accumulate, and the Left image undergoes double warping.
- **Center-Anchored**: We treat the Center image as the canonical view (I).

$$H_{L\to C} : \text{Warps Left to Center} \tag{2}$$

$$H_{R\to C} : \text{Warps Right to Center} \tag{3}$$

2.6 Seamless Blending and Intensity Correction

To overcome intensity variations and visible seams, we implemented two enhancements:

1. **Exposure Compensation**: We compute the mean brightness in the overlapping regions ($R_{overlap} = I_1 \cap I_2$). A global gain factor g is computed to equalize the intensity of the side images to the center anchor:

$$g = \frac{\mu(I_{center})}{\mu(I_{side})}, \quad I'_{side} = g \cdot I_{side} \tag{4}$$

2. **Distance Transform Blending (Feathering)**: Instead of a hard cut or simple averaging, we assign weights $w(x, y)$ to each pixel based on its Euclidean distance from the nearest image edge using a Distance Transform. The final pixel value is a weighted average:

$$P(x, y) = \frac{\sum w_i(x, y) I_i(x, y)}{\sum w_i(x, y)} \tag{5}$$

This ensures smooth transitions and eliminates visible stitching artifacts.

3 Experimental Results

3.1 Naive Stitching Analysis

We first attempted a naive concatenation of the images (resized to equal height).

Figure 2: Result of Naive Stitching. Note the visible seams and vertical misalignment (circled).

As seen in Figure 2, this approach fails significantly because:

1. It assumes pure camera translation (trucking) rather than rotation (panning).
2. It ignores perspective distortion.
3. It cannot handle vertical misalignment.

Parameter	**Value**
Image dimensions (resized)	6048×8064 pixels
Overlap percentage	35%
Overlap width (L-C and C-R)	2,116 pixels
Blending method	Simple averaging (50/50)
Final canvas dimensions	13912×8064 pixels

Table 1: Quantitative parameters for naive stitching approach.

The naive approach uses direct overlay with 35% overlap and simple averaging in overlapping regions. While computationally efficient, this method produces visible artifacts because it does not account for geometric transformations between views.

3.2 Quantitative Analysis

We provide detailed statistics on keypoint detection, matching, and inlier ratios (Table 2).

Image Pair	Keypoints (Img1)	Matches	Inliers	Inlier %
Left-Center	387,635	22,718	14,890	65.5%
Right-Center	121,839	11,680	8,266	70.8%

Table 2: SIFT keypoints detected, feature matches after Lowe's test, and RANSAC inliers.

The high keypoint counts (100K–400K per image) reflect the rich texture in the scene. The inlier percentages (65–70%) demonstrate that RANSAC successfully filters outliers, retaining only geometrically consistent correspondences.

Figure 3: Top 50 feature matches between Left and Center images (before RANSAC filtering).

Figure 3 visualizes the SIFT feature correspondences. Even with some outliers visible, the majority of matches correctly identify corresponding features, enabling robust homography estimation.

3.3 Robust Stitching Analysis

Our final method using SIFT, RANSAC, and Center-Anchored warping produces a seamless panorama. Figure 4 demonstrates:

- **Perspective Correction**: Straight lines are preserved across boundaries.
- **Alignment**: Overlapping regions match perfectly due to robust homography estimation.
- **Field of View**: The center-anchored approach preserves the wide-angle view without excessive distortion at the edges.

4 Future Work

While our implementation produces visually acceptable results, several enhancements could further improve quality:

- **Multi-band Blending**: Using Laplacian Pyramids to blend low frequencies and high frequencies separately, preventing ghosting while maintaining seamless color transitions.
- **Graph Cut Seam Finding**: Finding the optimal seam path that minimizes intensity differences, rather than simple distance-based mixing.
- **Bundle Adjustment**: Global optimization of all camera parameters simultaneously.
- **Cylindrical/Spherical Projection**: For full 360° panoramas to reduce distortion.

5 Conclusion

This assignment validated the effectiveness of invariant features for automatic image stitching. While simple concatenation is computationally cheap, it is geometrically invalid for panorama generation. The combination of SIFT features for matching and RANSAC for robust estimation allows for high-quality stitching even in the presence of wide baselines and clutter.

Figure 4: Final Panorama using SIFT + RANSAC + Center-Anchored Stitching.