DA526: Image Processing with Machine Learning

Project Report

Beyond the Horizon: Enhanced Panoramas

Group 6

234101002: Abhishek Verma
234101005: Ajaypal Singh
234101013: Dhruvkumar Babubhai Kakadiya
234101019: Jash Vipul Ratanghayra
234101024: Kishan Hitendra Thakkar

1. Problem Statement

Photo stitching, also referred to as image stitching, involves merging multiple photographic images that share overlapping views to form a panorama or a high-resolution image.

Successful image stitching demands precise overlaps and consistent exposures across images to achieve seamless integration. The process involves employing computer algorithms to seamlessly combine images into a larger composition, without visible seams. These algorithms are integral to features like "image stabilization" in camcorders and are used to generate the high-resolution mosaics necessary for digital mapping and satellite imagery. Additionally, most modern digital cameras include these algorithms, enabling users to produce expansive, ultra wide-angle panoramas.

The project aims to study and implement the algorithms for image stitching and survey the challenges faced while generating the panoramic image.

2. Dataset

For the dataset section of our project, we systematically collected a diverse set of photographic images across various locations within the IIT Guwahati campus using the cameras on our smartphones. The images were captured with a focus on maintaining a sufficient overlap between consecutive photographs to facilitate seamless stitching in the panorama creation process. To enhance the versatility of our dataset, we captured images intended for both horizontal and vertical panorama stitching, exploring the different orientations in which panoramas are commonly assembled.

Additionally, we frequently took more than two images of a single scene to explore and refine the stitching of multiple images into a cohesive panorama. This approach allowed us to test the algorithm's capability to handle extended sequences of photos, a critical factor for creating wide-ranging panoramas.

To ensure the data was compatible with our processing requirements, we cropped the images to a standardized size and scaled down the resolution. This modification was intended to speed up the processing time without significantly compromising the quality necessary for effective image stitching. The dataset includes images taken under varied lighting conditions to test the robustness of our stitching algorithm. We captured scenes during different times of the day, including daylight and nighttime settings, to assess how changes in ambient light affect the stitching process.

Furthermore, we made an effort to include images with minimalistic features, such as clear skies with few or no clouds, to evaluate the algorithm’s performance in scenarios with less texture and distinct landmarks. This diverse collection of images from our campus not only provides a comprehensive base for testing our image stitching algorithm but also simulates real-world conditions that users might encounter.

Dataset image samples:

Figure: Captured images at night time, shows left and right side views, taken continuously

Figure: Captured images of clouds, shows left and right side views

Figure: Captured images of nanotech dept. with the aim to stitch multiple views

3. Related Work

In recent years, significant advancements have been made in the field of panoramic image stitching, with researchers exploring various techniques to address the challenges posed by three-dimensional, rotational images with illumination variation. Bind et al. in his paper "A Robust Technique for Feature-based Image Mosaicing using Image Fusion, 2013" introduced a panoramic stitching approach tailored for such scenarios, leveraging the robustness of Scale-Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF) algorithms. SIFT, known for its scale and rotational invariance as well as noise robustness, and SURF, offering similar properties along with illumination invariance and computational efficiency, were employed for initial image alignment.

Furthermore, Bind et al. utilized Discrete Wavelet Transform (DWT) for seamless blending of the stitched images. This combination of feature-based alignment and wavelet-based blending proved effective in producing high-quality panoramic outputs from input images with varying illumination conditions and perspectives.

In a similar vein, Antony and Surendran in their paper "Satellite image registration and image stitching Satellite image registration and image stitching, 2013" developed a stitching technique specifically tailored for satellite imagery, emphasizing geometric alignment and seamless blending. Their approach involved initial image registration to geometrically align the input images, followed by a stitching algorithm utilizing the alignment estimations to blend the images seamlessly. Notably, their system demonstrated versatility by supporting images in various formats, including JPEG, TIFF, GIF, and PNG.

However, Antony and Surendran's method faced challenges when dealing with images exhibiting significant differences in lighting conditions. To mitigate this issue, they recommended normalizing the images prior to applying the stitching method, thus enhancing the overall performance and ensuring accurate alignment and blending

These pioneering works highlight the evolving landscape of panoramic image stitching, showcasing innovative approaches to address complex challenges such as illumination variation and geometric alignment, thereby paving the way for further advancements in this domain.

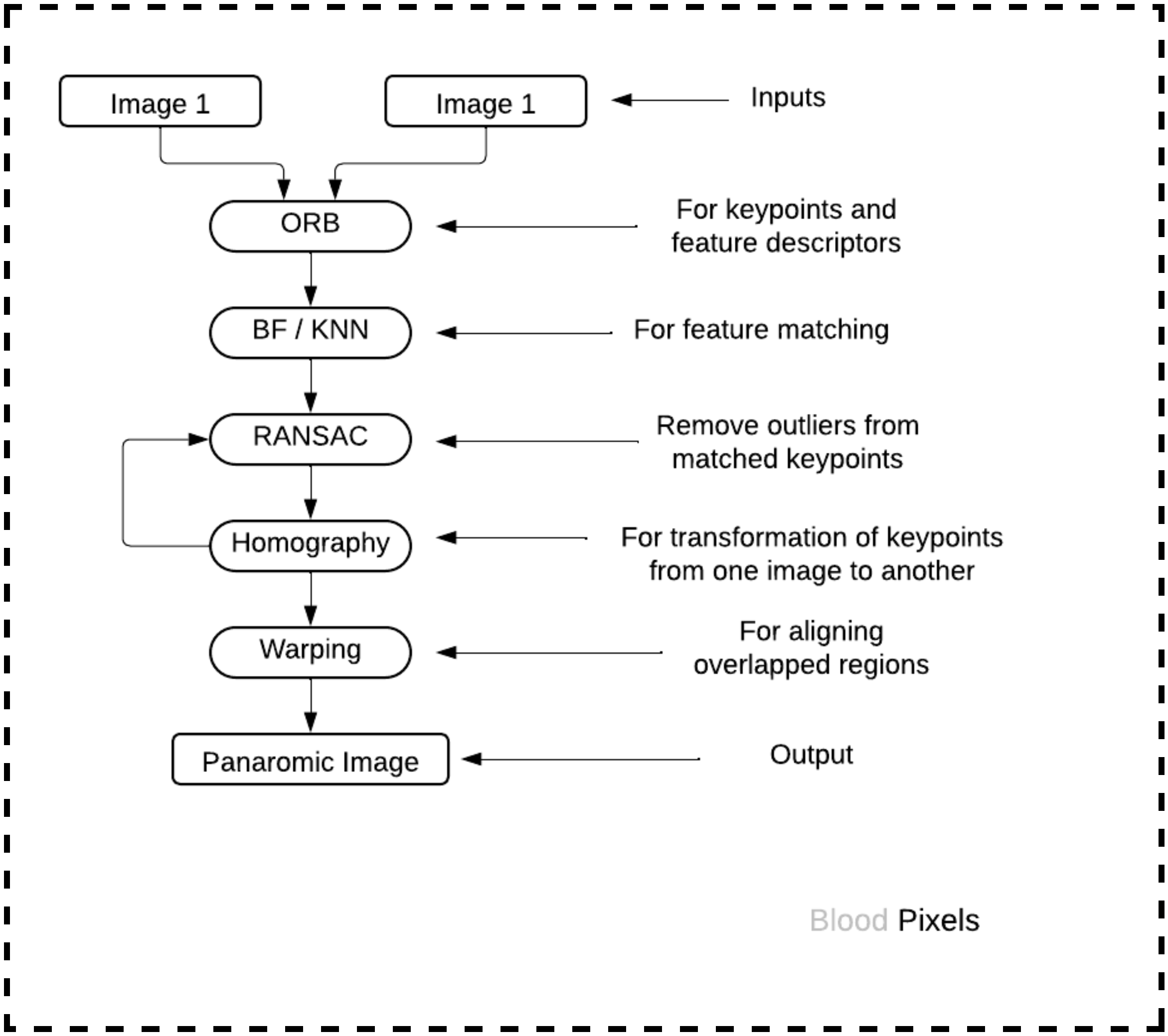
In further exploration of feature-based detectors for image stitching, Adel et al. in his paper "Real time image mosaicing system based on feature extraction techniques, 2014" conducted a comparative study evaluating several prominent techniques. Their investigation encompassed the Harris corner detector, GoodFeaturesToTrack detector, SIFT, SURF, FAST, MSER detector, and ORB technique, assessing factors such as detection rate, computational time, and accuracy.

Their experimental findings highlighted SIFT as a robust algorithm, albeit with relatively higher computational demands. On the other hand, both ORB and MSER algorithms exhibited robustness comparable to SIFT, with ORB emerging as the fastest technique. Moreover, Adel et al. introduced a real-time image stitching system based on the ORB feature-based technique, further demonstrating its efficacy.

Subsequent experiments comparing ORB to SIFT and SURF reaffirmed ORB's superiority in terms of speed, performance, and minimal memory requirements.

These findings contribute to the understanding of feature-based detectors in image stitching applications, showcasing ORB as a promising alternative with real-time capabilities and high performance.

4. Method:



- **Identify distinctive features in each image using a feature detection algorithm (e.g., SIFT, SURF, ORB).**

1. **Keypoint Detection:**

   FAST Keypoint Detector: ORB starts with the FAST (Features from Accelerated Segment Test) algorithm to detect corners. FAST is known for its high speed, which comes from only examining a circle of sixteen pixels around each candidate point to determine if it qualifies as a corner based on a set brightness test.
   Scale Invariance: To add scale invariance, ORB uses a pyramid approach, where the original image is repeatedly scaled down (creating a "pyramid" of scaled images), and FAST is applied at each level of this pyramid. This approach helps in detecting keypoints at different scales.

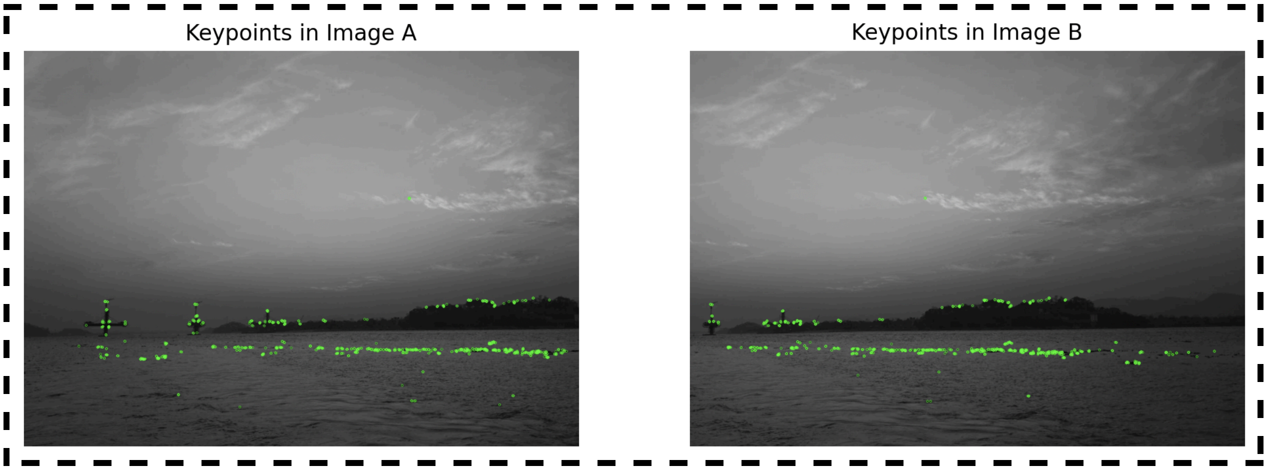


Figure. Images with highlighted feature descriptors

2. **Descriptor Generation:**

Rotated BRIEF: After determining the orientation, ORB uses the BRIEF (Binary Robust Independent Elementary Features) descriptor, but with a key modification. BRIEF normally creates a binary string out of intensity differences between pairs of pixels in a smoothed image patch around each keypoint. ORB rotates the BRIEF descriptors based on the keypoint orientation, making the descriptor rotation invariant.

Binary Descriptor: The result is a binary string that efficiently summarizes the local gradients and patterns around a keypoint. These descriptors are robust to noise and are very quick to compute and compare, which is ideal for real-time applications.

- **Feature Matching using Brute Force and K Nearest Neighbour.**

Principle: The k-NN algorithm identifies the closest training examples in the feature space. In the context of feature matching, it compares the feature descriptors of keypoints from two images to find matches based on their distance in the descriptor space.

Distance Metric: The Hamming distance is used when dealing with binary descriptors such as those produced by ORB. This metric effectively measures the similarity between two binary strings.

Efficiency: While k-NN can be computationally intensive, its efficiency can be enhanced by using optimized data structures such as KD-trees or Ball Trees for quicker nearest-neighbor searches.

Robustness: k-NN matching allows selecting the best match by considering multiple nearest neighbors, rather than just the closest one. This method helps in reducing the effects of noise and outliers in the matching process.

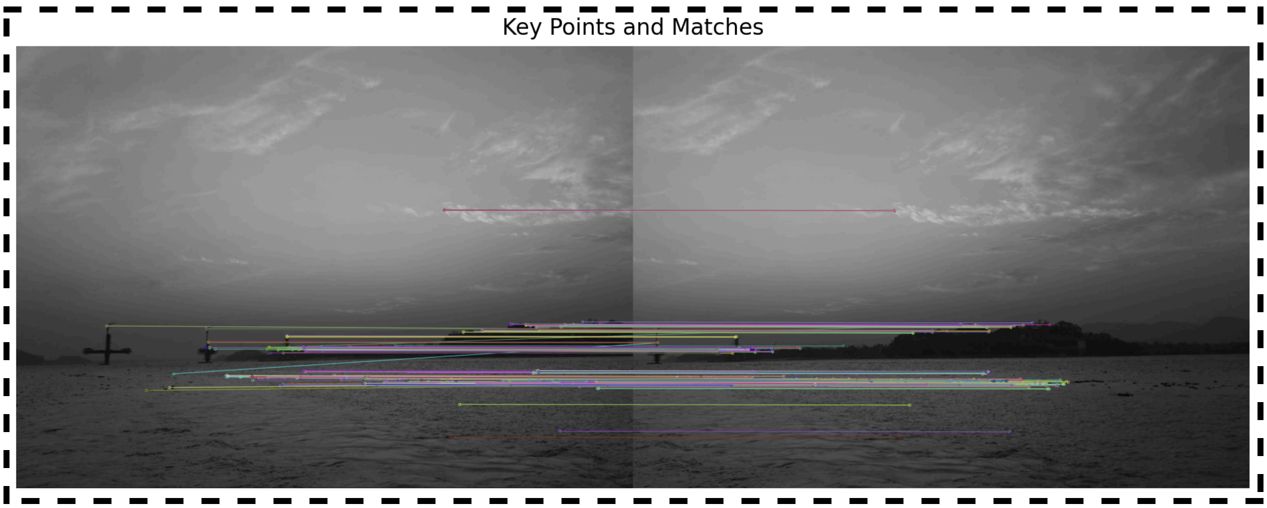
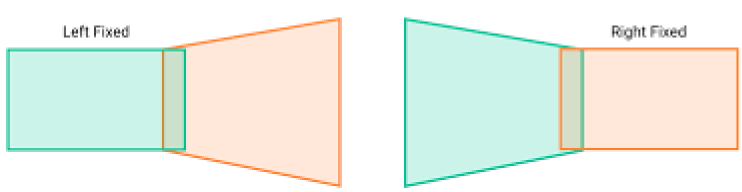


Figure. Shows the matches between key points of the overlapped features

- **Homography and Ransac Technique for projection.**

Random Sample Consensus algorithm — more popularly known by the acronym RANSAC. This is an iterative and a non-deterministic algorithm that helps in eliminating outliers. This algorithm is commonly used to solve computer vision challenges.



1. Randomly select a smaller set of points (n) from the entire distribution (N). Use least squares regression to determine the linear equation which fits the n points.
2. Determine the average of the distance of every point N from this line. This score can be considered to be a measure of the goodness of the line.
3. Keep track of the score. If this score is less than the best score obtained from previous iterations then discard the older linear equation and select the current linear equation.
4. Go back to the first step and continue iterating till you have completed a predetermined number of iterations.
5. Stop the algorithm when a predetermined number of iterations have been completed. The linear equation available at the end of the iterations is possibly the best candidate line
   We can see that the algorithm is not deterministic and hence the name Random in the acronym RANSAC. It is possible that you may not get the best model.

   Algorithms pseudo Code:

1. MAX = max iterations.
2. N = number of points to pick in every iteration. Could be initialized to 2.
3. best_model = equation of the line with best_error . Initialize to NULL.
4. best_error= The lowest error (average distance) obtained so far. Initialize to a large number threshold_error=if the distance of a point from a line is below this value then the point is classified as an inlier otherwise outlier.
5. Threshold_inliers = minimum number of inliers for a model to be selected.
6. K = count of iterations completed. Initialize to 0.

- **Warping:**

1. **Mapping Function**: This process is driven by a mapping function, which defines how pixels in the source image are relocated in the target image. The nature of the mapping can be affine (linear transformation including rotation, scaling, translation) or more complex (non-linear, such as projective and polynomial transformations).
2. **Interpolation:** Since the mapping can result in pixel coordinates that are not aligned with the discrete grid of the output image, interpolation is used to compute pixel values at these new locations. Common interpolation methods include nearest-neighbor, bilinear, and bicubic interpolation.

5. Experiments and Results:

Figure. Panorama of the nanotech department.

Figure. Panorama of Brahmaputra river with sky

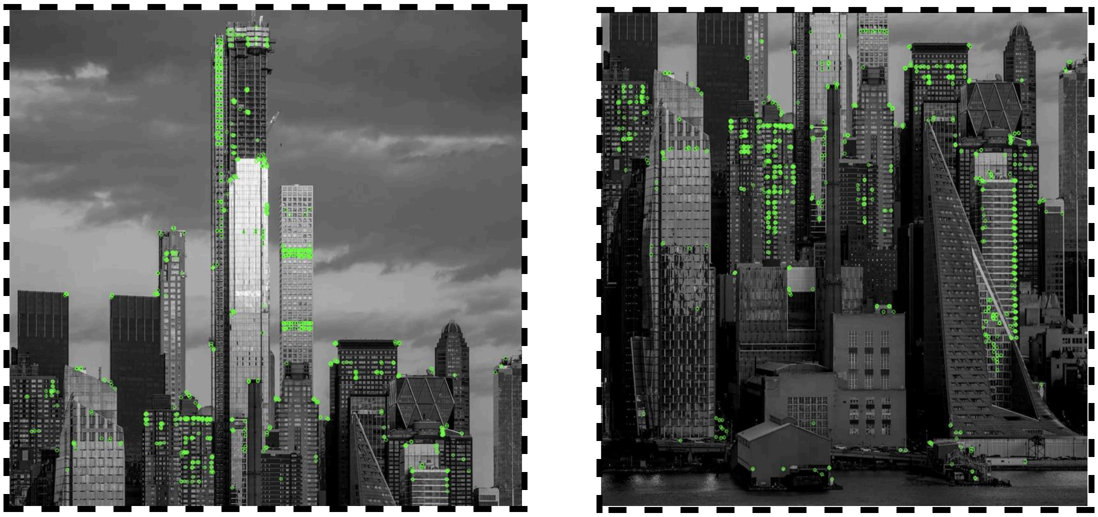


Figure. Highlighted features of building with aim to stitch vertically

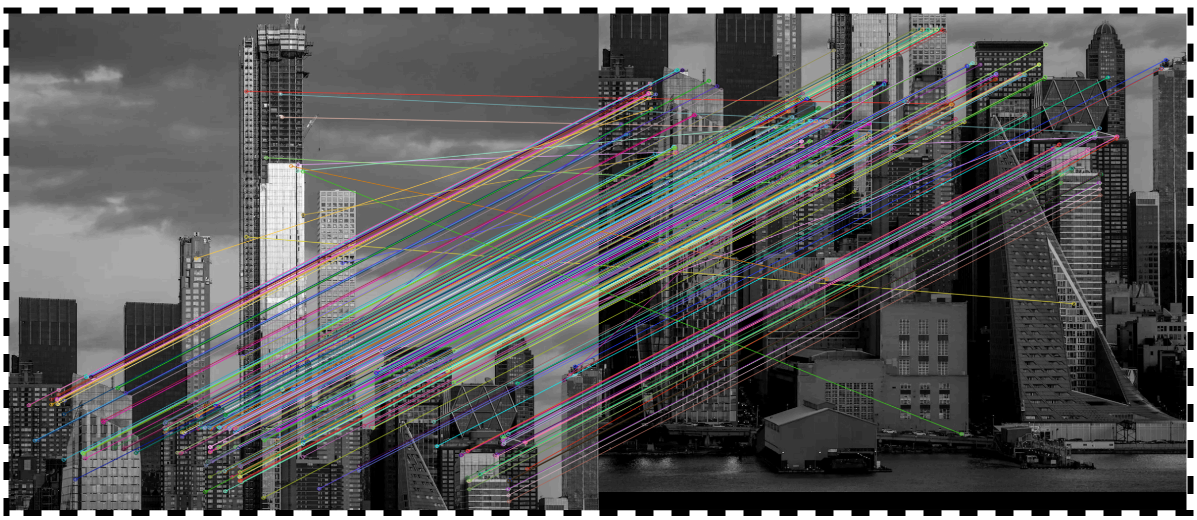
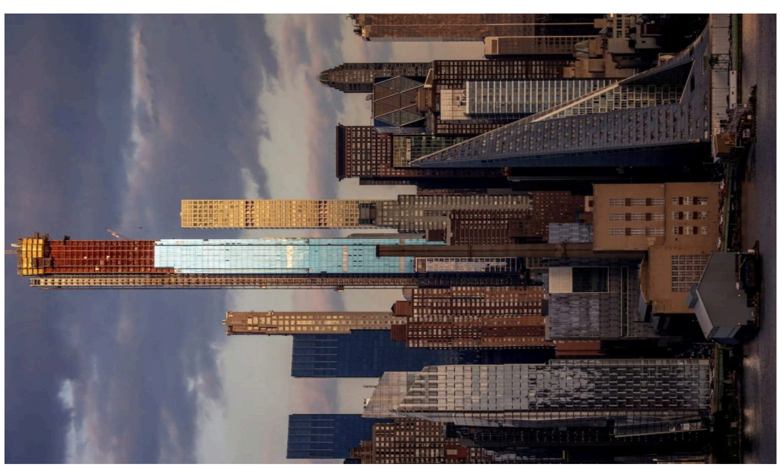


Figure. Matched features of the images of buildings



6. Conclusion

Throughout this project, we conducted extensive experiments on panorama image stitching using a dataset collected from various locations within the IIT Guwahati campus. Our efforts encompassed stitching images both horizontally and vertically, incorporating multiple images to create extensive panoramas, and modifying environmental variables such as lighting conditions. While we successfully generated panoramic images as intended, several challenges and limitations were observed, highlighting areas for future research and development.

Challenges and Limitations Encountered:

1. Motion-Induced Inconsistencies: Our experiments revealed that even slight motions during image capture could lead to misalignments and distortions in the stitched panoramas. These inconsistencies were particularly evident in scenarios involving multiple image stitching where maintaining absolute camera steadiness was more challenging.

2. Brightness and Contrast Disparity: Variations in brightness and contrast between the stitched images were visible, affecting the visual continuity of the panoramas. These disparities were most pronounced when combining images taken under different lighting conditions.

3. Optimal Threshold Estimation: Determining the optimal threshold for image feature detection for instance, stitching images with fewer features, such as skies, required a lower threshold for feature detection to ensure sufficient data points for alignment.

4. Occurrence of Blank Spaces: In some cases, especially in multi-image stitching, blank spaces appeared in the final panorama. These were primarily due to inadequate overlap or incorrect alignment, emphasizing the need for precise control over image capture. These blank spaces were eliminated using cropping techniques.

5. Blending of Images: The blending of images did not achieve the desired level of seamlessness. This was apparent where different textures and colors met, indicating a need for more advanced blending algorithms that can more effectively mitigate visible seams.

Future Scope

To address the above challenges, future work could focus on the following areas:

- Enhancing Motion Stabilization Techniques: Developing more sophisticated image stabilization techniques could mitigate the effects of motion-induced inconsistencies, particularly in dynamic environments.

- Adaptive Brightness and Contrast Adjustment: Implementing adaptive algorithms that automatically adjust brightness and contrast during the stitching process could improve the visual consistency of panoramas.

- Advanced Blending Algorithms: Research into more complex blending algorithms that consider both color and texture information might result in more seamless panoramas, especially in complex overlap regions.

By addressing these limitations and continuing to innovate in the fields of image processing and computer vision, with potential applications expanding into including Virtual Reality (VR) where it can enhance immersive experiences. It's also pivotal in Real Estate and Interior Design, offering detailed virtual tours. Autonomous Vehicles can benefit from improved spatial awareness and visualization. Remote Surveillance, Tourism, and Travel industries can leverage panoramic images for better monitoring and enhanced visitor experiences. Cultural Heritage Documentation can preserve sites in digital detail.

References

A survey on image and video stitching
https://www.sciencedirect.com/science/article/pii/S2096579619300063

A Review Over Panoramic Image Stitching Techniques
https://iopscience.iop.org/article/10.1088/1742-6596/1999/1/012115

Image Stitching System Based on ORB Feature Based Technique and Compensation Blending
https://pdfs.semanticscholar.org/cf0d/3838b87c0f14f5941f78252c444126ae36bc.pdf

Github
https://github.com/DZ521111/Beyond-Horizon/