Assignment

Image Stitching

Members

Arya Dwivedi (Enrolment No. 18/11/EC/045)

Sanya Agrawal (Enrolment No. 18/11/EE/023)

Gargi Singh (Enrolment No. 18/11/EC/030)

Deepanshu Jain (Enrolment No. 18/11/EC/047)

Naveen Kumar

1. Steps we followed:

1) Preprocessing
2) Dominant Frame Selection
3) Feature Matching
4) Calibration
5) Blending

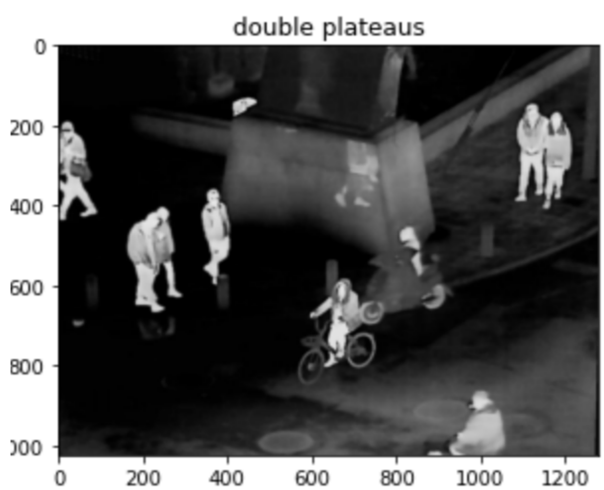
2. Datasets:

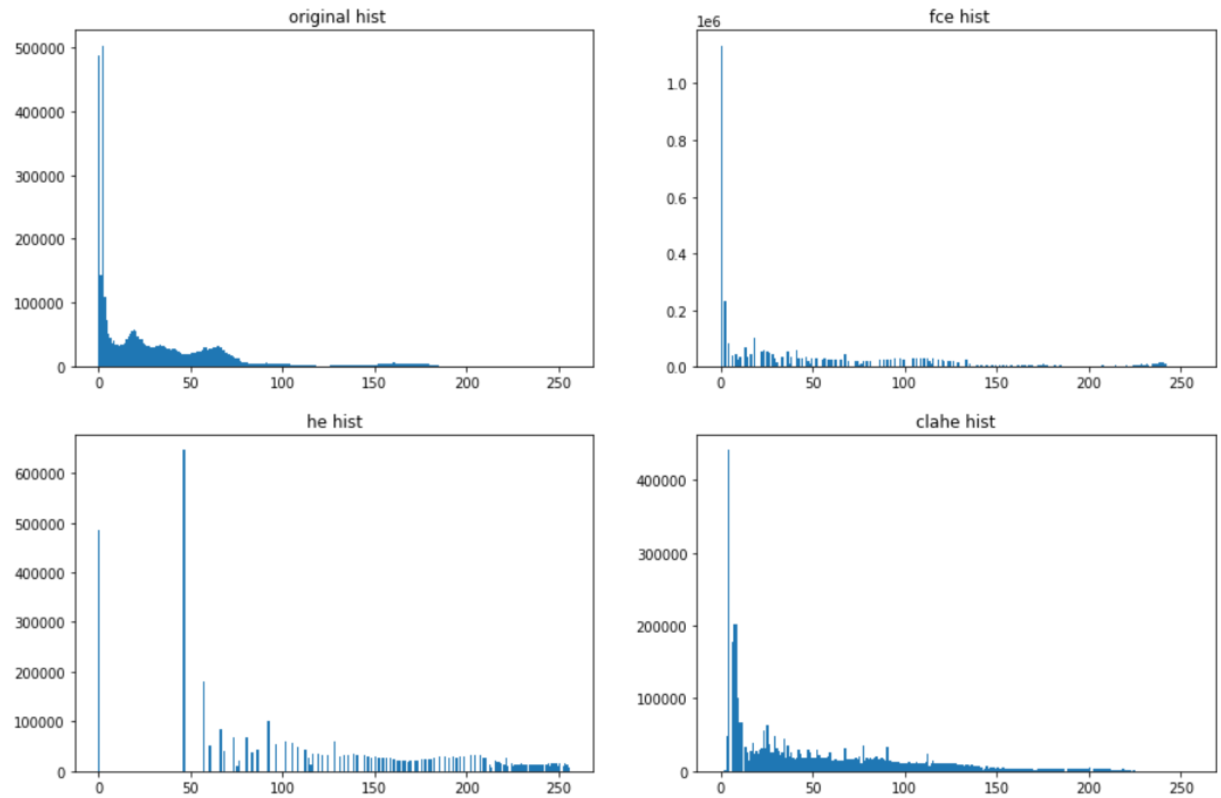
2.1. LLVIP
2.2. Project_IR
2.3. FLIR_ADAS_v2

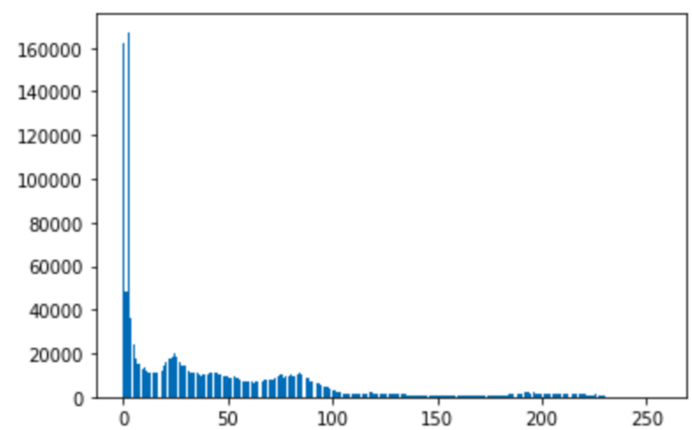
3. Preprocessing Techniques used:

Raw infrared images suffer from poor contrast and low resolution when compared to visual images, making it difficult to distinguish the target from the background. This limits the performance of Image Stitching.
Fuzzy-logic enhancement, Contrast Limited Adaptive Histogram Equalization (CLAHE) and Double Plateaus histogram equalization have been implemented based on published papers. These were each compared to the classical technique of histogram equalization.

Most preprocessing techniques ended up amplifying background noise while suppressing detailed information.
What works best is the double plateau which constrains the background noise using an upper threshold value while simultaneously, a lower threshold protects the detail information from getting suppressed.
A comparison between different methods is shown below:-








```
Text(0.5, 1.0, 'double plateaus hist')
```

double plateaus hist

4. **Extraction of Dominant Frames**:

The aim of dominant frame selection is to find the dominant frame set of the image series, in which the adjacent frame pair has a balanced overlap ratio not only providing enough matched points to image registration, but

also skipping too much repetitive content. It helps us select our query image and train image to stitch from a vast dataset (which was not very relevant in our case because we did not have an expansive dataset but it was still helpful). Therefore, the computation cost of feature extraction and feature registration could be reduced, which contributes to the efficient stitching of the infrared image series.With the aforementioned steps, a threshold value τ is set to determine which frames are collected into the dominant set. When the similarity value is lower than the threshold τ, the frame is added to the dominant frame set, and the aforementioned steps are repeated until the last image frame is encountered. Finally, the dominant frame subset could be obtained to create the panoramic scene.

Algorithms used for feature detection and descriptors :

1. SIFT
2. MIFP using EOH descriptor
3. ORB
4. KAZE
5. BRISK

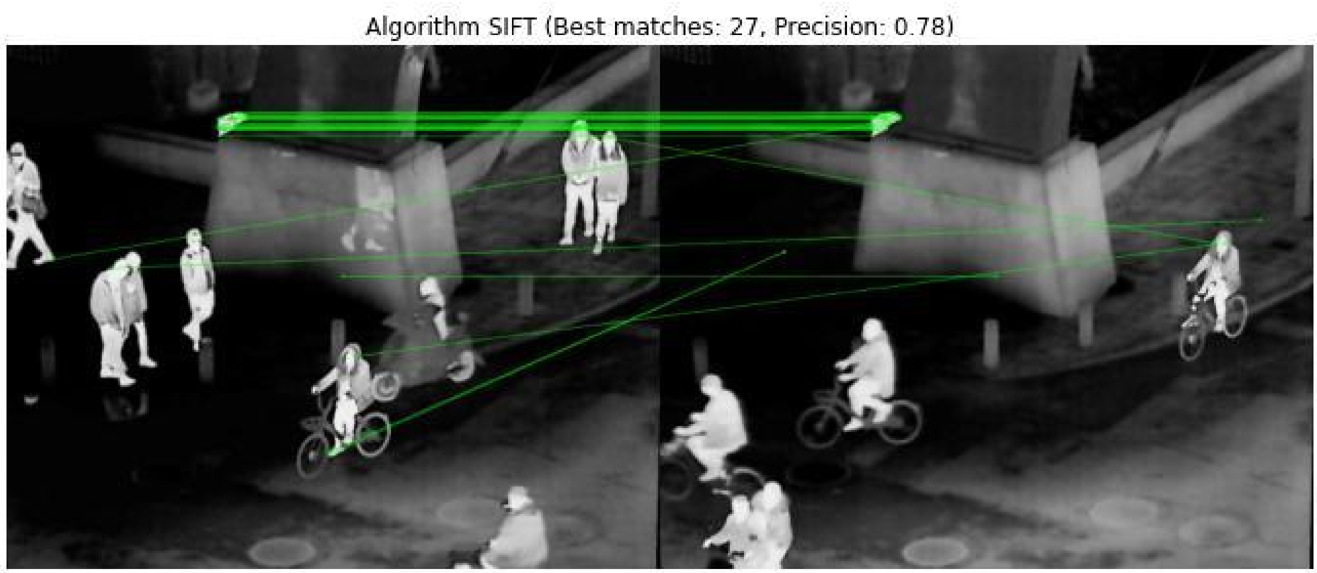
MIFP(multi spectral image feature points) using EOH works best

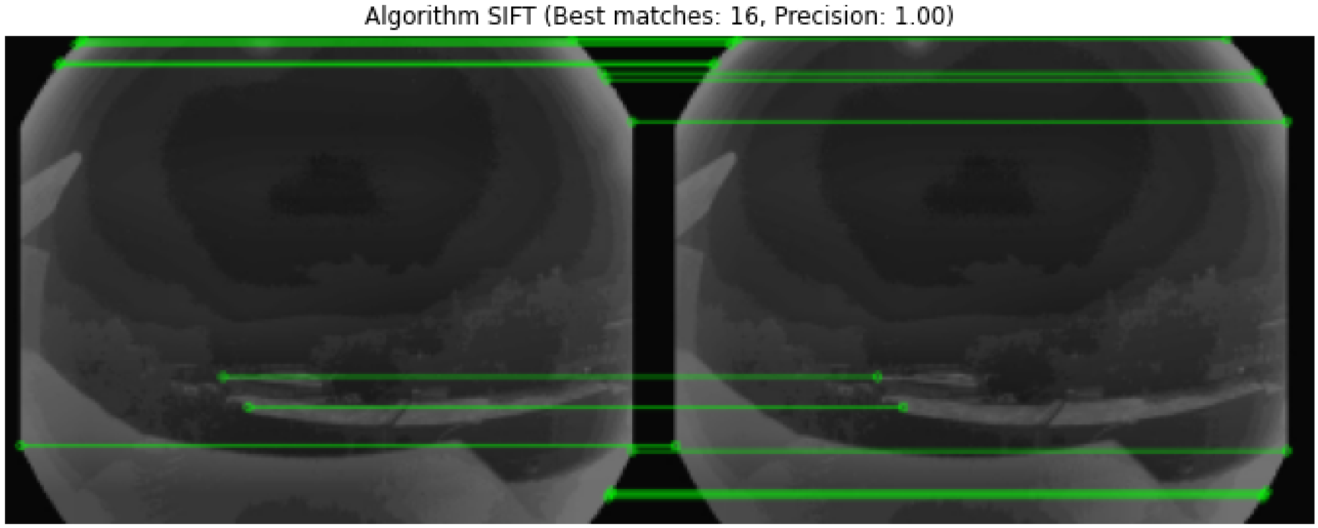
The proposed scheme consists of a scale-space pyramid, like the one used by SIFT. Similarly, invariant features are used, but by modifying the feature vector in such a way to incorporate spatial information from the contours of each keypoint without using gradient information. This allows us to generate a correlated parameter space in both the Infrared images. This edge orientation histogram describes the shapes and contours from IR images, keeping in the scale-space their invariance.

Recent studies have shown that as the spectral bands go away from visible spectrum, classical feature descriptors generally used for finding matching and registration of images in the visible spectrum are useless. This is also clearly observed in our results.

The difficulty in finding correspondences between feature points from Infrared images results from the nonlinear relationship between pixel intensities. Variations in IR intensities are related to variations in the temperature of the objects, while variations in VS intensities come from color objects and light reflections. Therefore, this nonlinear relationship results in a lack of correlation between their respective gradients. Furthermore, IR images appear smoother, with loss of detail and texture , so that the detection of corners, as candidates for local descriptor points, is also poorly favored. In conclusion, most of the image processing tools that use gradients of pixels based descriptors need to be adapted, or otherwise they become useless. Hence EOH which ses the edge points distribution of four directional edges and one non-directional edge to construct the feature description, works best.

SIFT(**scale-invariant feature transform**) :





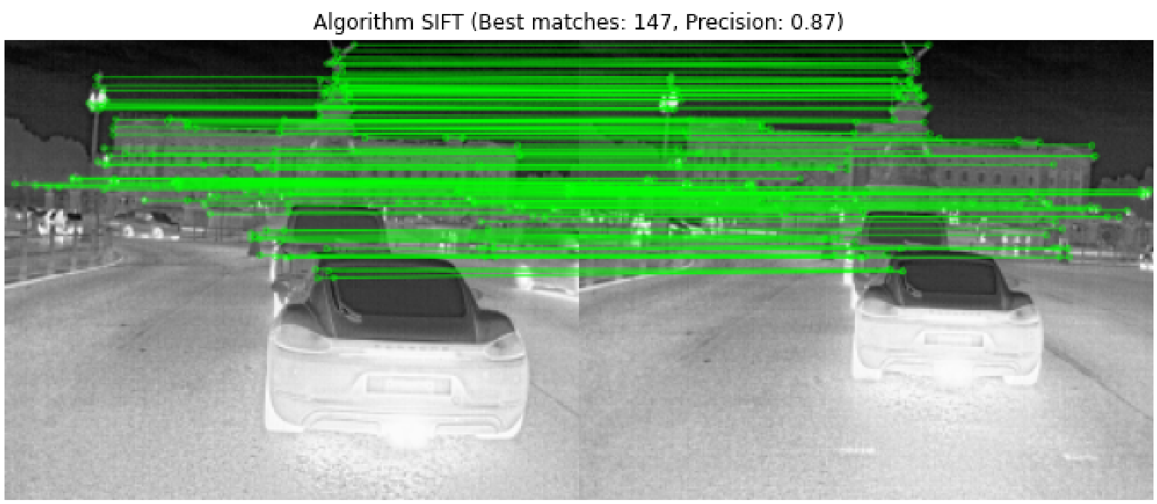
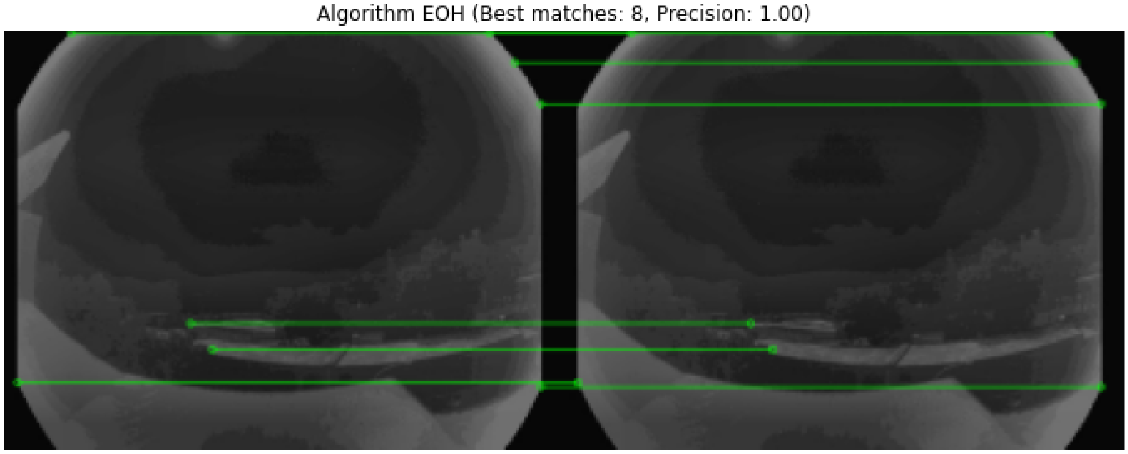


Fig. SIFT for keypoints, descriptors and KNN for Feature Mapping

MIFP using EOH descriptor (Edge Oriented histogram):

Algorithm EOH (Best matches: 15, Precision: 0.47)

Algorithm EOH (Best matches: 8, Precision: 1.00)



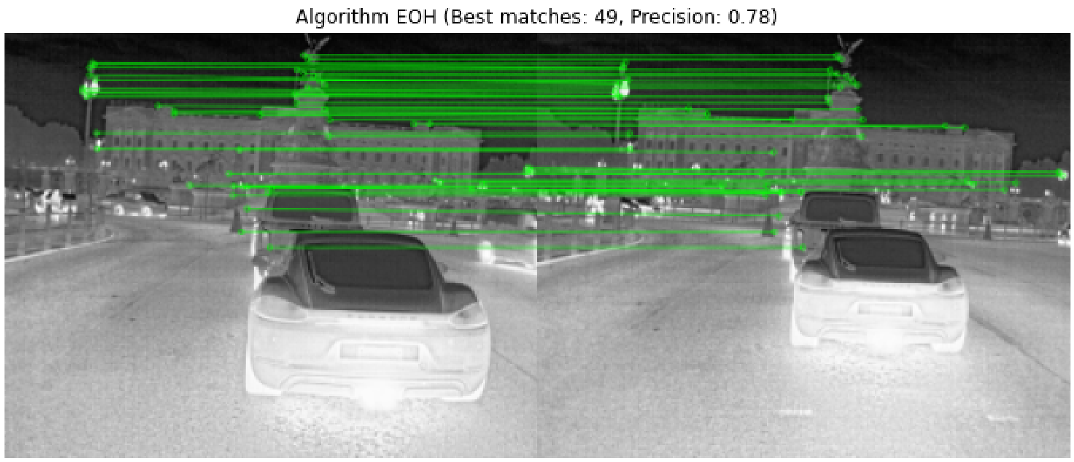


Fig. MIFP for keypoints, descriptors and KNN for Feature Mapping

ORB

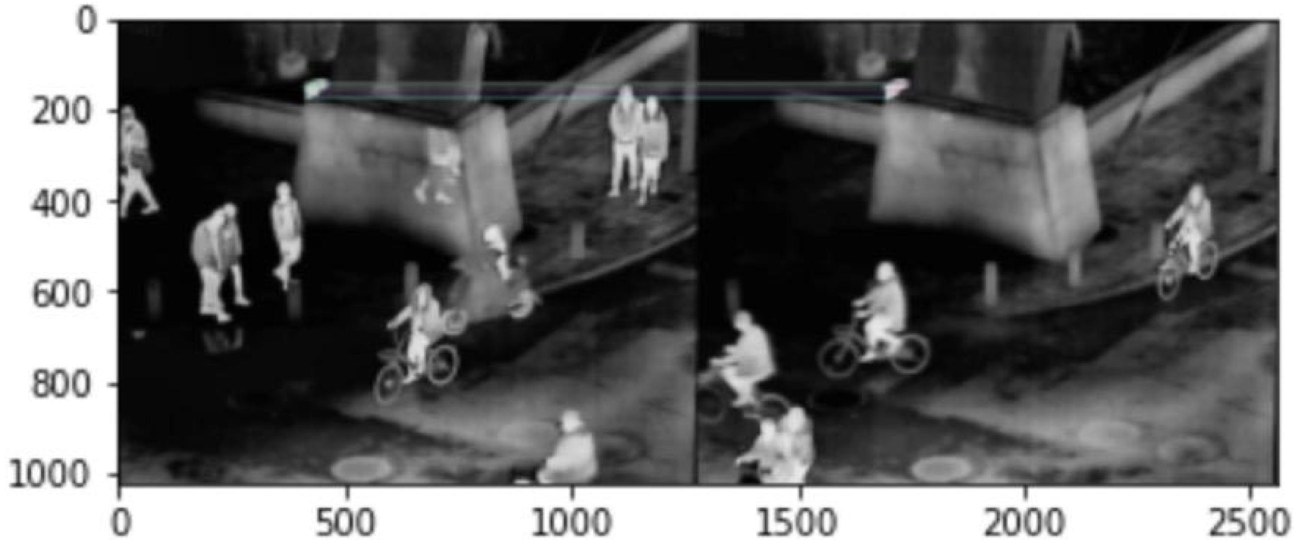


Fig. ORB for keypoints, descriptors and KNN for Feature Mapping

KAZE

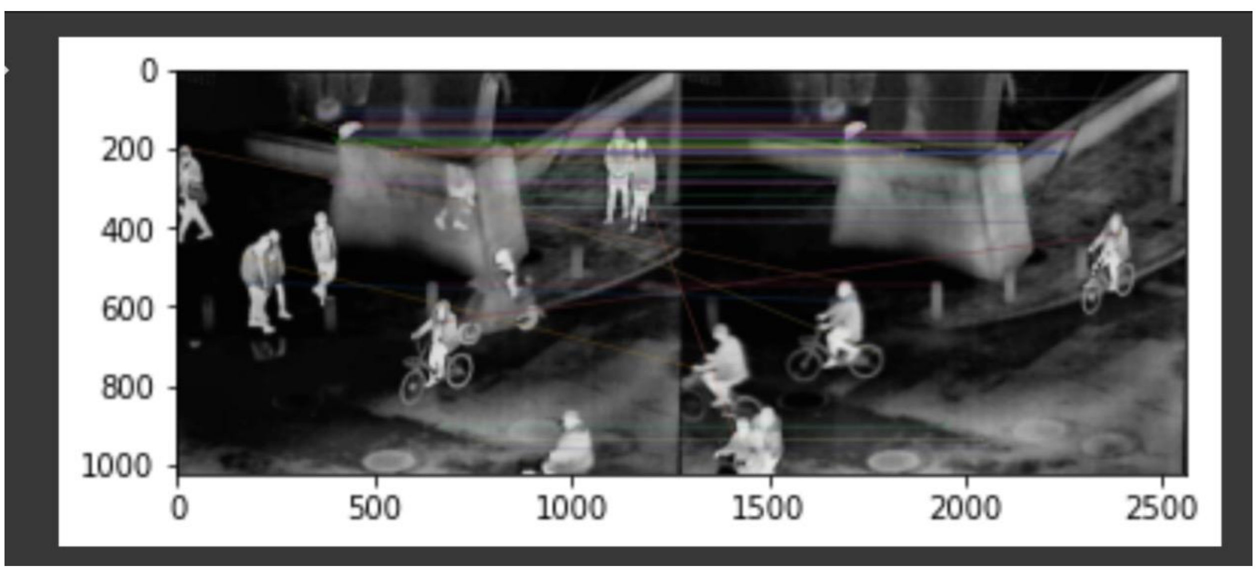


Fig. KAZE for keypoints, descriptors and KNN for Feature Mapping

BRISK

```
Using: knn feature matcher
Raw matches (knn): 1633
```

Fig. BRISK for keypoints, descriptors and KNN for Feature Mapping

Now, we need to take these points and find the transformation matrix that will stitch the 2 images together based on their matching points.
For that We have used Homography - It maps points from one plane (image) to another. We used RANSAC as our model is sensitive to outliers,
RANSAC solves this problem by estimating parameters only using a subset of inliers in the data.

6. Blending

Fig. Perspective Blending

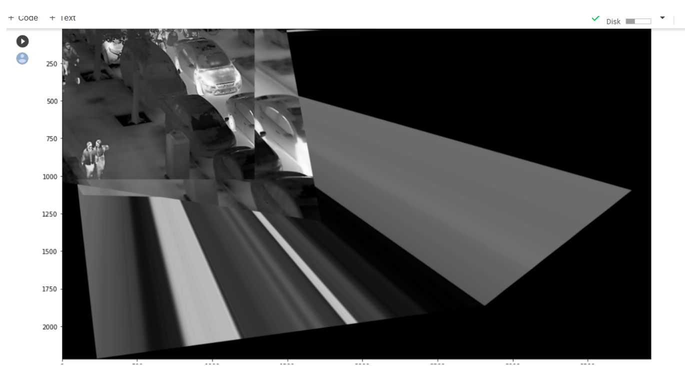
Fig. Linear Blending

We have used perspective transformation to one of the images. Basically, a perspective transform may combine one or more operations like rotation, scale, translation, or shear. The idea is to transform one of the images so that both images merge as one.

When we tried blending the images we were unable to get good results even after all the preprocessing.

Reasoning for poor results : The images did not have any movement along the x-axis specially for the images in the LLVIP and FLIR_ADAS_V2 dataset.
Since objects existing in 3D are projected onto a 2D image using homogeneous coordinates in the image, all are judged at the same depth.
That is, the moment the change is made from 3D to 2D, depth (z-axis) information disappears. Because of this problem, if matching is performed based on the features existing in the original area, it becomes difficult to match the features . These images require some kind of depth stitching (probably along the z-axis) that we were unable to do using the existing techniques that we knew of.
For example: in the following example the stitching results are poor because there is an incorrect correspondence point. Also, since several corresponding Points are projected in two dimensions, and three-dimensional depth information is lost.To minimize this problem, the RANSAC technique is used, but it is still difficult to obtain a perfect stitched image since these techniques also have limitations specially when it comes to multispectral images.

To overcome this a more accurate feature description method is required.



7. Conclusion:

Double Plateau Algorithm gave the best results for presprocessing. Fuzzy Contrast also gave us comparable outputs.

A scale-space pyramid, similar to the one employed by SIFT, is used in the proposed approach. Invariant features are utilised in a similar fashion, but this time by changing the feature vector to include spatial information from the contours of each keypoint without employing gradient information. In both the VS and LWIR pictures, this allows us to build a correlated parameter space. A descriptor based on the edge histogram is used in our proposal.

EOH gave us the best results as feature descriptor since EOH has a better matching ability than SIFT (scale-invariant feature transform) on multispectral images.

For feature matching, KNN produced better results as compared to brute force. Instead of returning the single best match for a given feature, KNN returns the k best matches.

8. Future Scope

We also tried to implement Panorama Stitching based on Asymmetric Bidirectional Optical Flow, to accurately stitch overlapping areas where the motion field of overlapping areas is calculated using an optical flow algorithm, and a more accurate stitching image is generated with the new warping map. but we were unable to implement the complete code due to a few errors.

9. References:

1. https://towardsdatascience.com/image-panorama-stitching-with-opencv-2402bde6b46c
2. https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1706&context=theses
3. https://www.kaggle.com/code/scratchpad/notebookf4ce98a4ea/edit
4. https://www.researchgate.net/figure/Proposed-EOH-based-keypoint-image-descriptor-for-VS-LWIR-images_fig5_232742480
5. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11002/110022H/Algorithm-for-image-stitching-in-the-infrared/10.1117/12.2519537.short?SSO=1
6. https://www.geeksforgeeks.org/clahe-histogram-eqalization-opencv/ , https://docs.opencv.org/3.4/d6/db6/classcv_1_1CLAHE.html
7. https://opencv24-python-tutorials.readthedocs.io/en/latest/
8. https://journals.sagepub.com/doi/10.1177/1847979020980928