Journal of Al-Qadisiyah for Computer Science and Mathematics VOL.17.(3) 2025,pp.Comp 179–190

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Integrating Harris Corner Detector and SIFT Descriptor for High-Accuracy Feature Matching in Image Processing

***Ali Salam Al-jaberi* [1]*, Sura Fadhil Rahman* [2]*, Ihsan Faisal Raheem* [3]**

[1] *College of Computer Sciences and Information Technology, University of Al-Qadisiyah, Al-Qadisiyah, Iraq. Email:ali.s.aljaberi@qu.edu.iq*

2 *Computer Techniques Engineering, Imam AL-Kadhum College, Al-Qadisiyah, Iraq.Email: surafadhilrahman@gmail.com*

3 *Nizam College, Osmania University , Al-Qadisiyah , Iraq. Email: ihsanfaisal500@gmail.com*

ARTICLEINFO

Article history:
Received: 10/07/2025
Revised form: 06/08/2025
Accepted: 11 /08/2025
Available online: 30 /09/2025

Keywords:

Harris Corner, SIFT Descriptor, Processing, Feature Matching.

ABSTRACT

Feature matching is one of the primary operations in image processing and computer vision. In this work, the limitations of the Harris corner detector—sensitivity to scale and noise—are highlighted and its performance enhanced through the integration of the SIFT descriptor that is invariant to rotation and scale. The aim is to make feature matching more accurate and robust under varied conditions. The approach that was suggested was implemented on the Ishtar Gate, the Egyptian Pyramids, and the Save Iraqi Culture Monument photos. The process involved keypoint detection with Harris, description with SIFT, and feature matching using the nearest neighbor distance ratio. Precise and accurate results were obtained, particularly with structured images like the Ishtar Gate (100% precision, 99% accuracy). The joint approach is strong in overcoming conventional detector limitations and improving reliability in image matching. There is potential for further research in extending the model to medical imaging and real-time applications.

https://doi.org/10.29304/jqcsm.2025.17.32455

1. Introduction

Matching local features is one of the main tasks in computer vision. Harris Corner Detector, together with SIFT Descriptor, poses a strong combination for performing local feature matching. Harris Corner Detector is one of the most famous algorithms for finding corners, defined as points where the intensity changes strongly in all directions. The corners are the best features for feature matching because they might be invariant to translation, rotation, and small changes in lighting. Once the key points have been identified by the Harris detector, SIFT can describe the local region around those points in an invariant way with respect to scale and rotation. Image comparison is a task that determines the geometric alignment of two images of the same scene taken from different angles at the same or

∗Corresponding author: Ali Salam Al-jaberi

Email addresses: ali.s.aljaberi@qu.edu.iq

Communicated by 'sub etitor'

different times by the same or different devices. This is a crucial image processing job that is frequently used in computer vision and pattern recognition [1].

Harris corner detector is a feature extraction tool as well as a popular point-of-interest detector for its strong rotational invariance and scale, but not a constant meter for transformation that is more complex, including lighting change, and image noise. The local autocorrelation function of a signal, which quantifies the local alterations in the signal as spots are slightly shifted in various directions, serves as the foundation for the Harris corner detector [2].

The sift descriptor is a local feature used for image processing. It is very stable in feature point detection and matching, scaling, and rotation, as well as an affine transformation. The sift feature is used to identify corresponding points between two consecutive images by describing key steps such as on-band detection, precise localization of key points, custom routing, and a point descriptor of the keys [3].

Feature detection is the key step for all local feature descriptor-based systems working on image matching. The most straightforward and commonly used method of keypoint detection is the Harris Corner Detector without the use of a local descriptor. Because the Harris corner detector pays little attention to the richness of discriminative information of these discovered key points, the key points it detects are not qualified in terms of repeatability and informativeness for big-rotation and scale-change images. To increase their distinctiveness, the important points that were identified ought to be encoded into a representative feature description using the local geometric information.

The Harris Corner algorithm has several constraints affecting its efficiency in real-time applications. Its real-time efficiency is greatly degraded by a high time complexity, especially when detecting angles. Secondly, applying this method to automatically sort images results in the distortion of the final mosaic image. The algorithm also generates a high number of spurious features, which results in low positioning accuracy. It does not possess variable scale handling ability and is highly sensitive to noise, and thus less robust under varying imaging conditions. Moreover, it is not good at the detection of infrared images, particularly with low contrast and fuzzy boundaries. Another significant challenge is the classification of magnetic signals in computer interfaces and the difficulty in recognition of offline handwriting or pattern classification tasks, wherein the content is not dependent on the author. Finally, the formulation of recording issues as total distance in visual and infrared imaging further complicates its use in advanced image processing tasks. This paper aims to enhance the efficiency of feature matching in images and provide high accuracy in image processing tasks. It seeks to calculate features that provide high repeatability and high-quality matching performance, ensuring high robustness in varying conditions. The objective also encompasses feature extraction that does not change with regard to scale and rotation, minimizing the impact of changes in lighting, and accelerating the process of discovering features. These improvements are intended to boost the reliability and effectiveness of feature matching in practical applications.

2. RELATED WORK

In 2019, Bojanić et al. conducted a comprehensive benchmark to compare traditional handcrafted methods such as SIFT, SURF, ORB, FAST, BRISK, HARRIS, and others against deep learning-based models like LF-Net and super point. The study evaluated these techniques on the HPSequences dataset under various geometric and illumination conditions, considering tasks such as keypoint verification, image matching, and keypoint retrieval. Surprisingly, results showed that certain classical detector-descriptor combinations still rival or even outperform pretrained deep models, especially when execution time is taken into account, with super point and ORB emerging as the fastest [4].

In 2019, J. Qin et al., a possible discussion may start by saying that the retrieval of encrypted images represents an important operation in computational computing when one mentions computing. It is, in essence, one of the modules required to create image databases. At the time of encoding, the image features selected are those that have been well studied. An approach is to first extract the features from the image and then apply some sort of filter to those features, since they have to be put into an image reactor. Therefore, the components that make an image are termed "vectors" if the image is in a binary form or a series of "words" if the image is in a numerical representation [5].

In 2019, Feng, J., et al. Stable picture feature points inside the image can be identified via the Harris corner detection algorithm. Nevertheless, a few issues, including high computational costs, poor positioning precision, and the algorithm's limited ability to recognize corner locations in the absence of feature descriptors, have severely restricted its use. This paper proposes a new approach for extracting visual features. The image is first double-screened, after which the corner position is determined using the Harris corner detection technique and refined down to the sub-pixel level using an iteration process. Lastly, the feature point information is represented via a rotation-invariant rapid extraction descriptor. These outcomes demonstrate how well the suggested approach addresses the shortcomings of the Harris algorithm by quickly and precisely extracting stable characteristics from an image. Its potential applications [6].

In 2021, Efe, U., et al. A novel method is suggested for matching images, which utilizes characteristics acquired by a pre-made deep neural network to attain highly encouraging outcomes. More specifically, the suggested method uses the already-trained VGG architecture as the feature extractor and does not call for more training meant to enhance matching. An approximate geometric transformation is computed by first warping using fundamental concepts from the psychology discipline, such as the Mental Rotation paradigm. These approximations are only derived from the dense matching of nearest neighbors at the VGG network's terminal layer, which matches the outputs of the matching images. After this first alignment, the same process is repeated between reference and aligned images hierarchically to achieve a successful localization and matching [7].

In 2021, D. Reddy Edla et al. Harris Corner Select Z for pipelines that this image matching did not witness advanced equivalence, despite its role in computer vision tasks. To apply to Harris Z for a pipeline extraction process matching the picture and key points in it [8].

In 2022, Bellavia et al. introduced HarrisZ+, an improved iteration of the HarrisZ corner detector. HarrisZ+ is specifically optimized to benefit from recent breakthroughs in other areas of the image matching pipeline. Beyond mere parameter tuning, the approach introduces more sophisticated selection criteria, which result in more keypoints with a higher number and more even distribution as well as improved discriminativeness and localization accuracy. Plugged into today's image matching pipeline, HarrisZ+ achieves state-of-the-art results on a variety of matching benchmarks, even in comparison to end-to-end deep learning-based approaches. These findings demonstrate continued potential in traditional hand-crafted image matching methodologies [9].

In 2022, Wang et al. focused on the performance of interest point detectors in heterologous image matching, particularly between optical and SAR images with varying resolutions. Their research compared detectors such as SAR-Harris, UND-Harris, Har-DoG, Harris-Laplace, and DoG, taking into account scale adaptability, nonlinear intensity differences, uniformity of distribution, and the accuracy of image alignment. The experiments showed that while SAR-Harris had greater adaptability to scaling variance, DoG had the best detection efficiency, and Har-DoG had the best final image alignment, providing useful insights on what detector is appropriate for specific remote sensing conditions [10].

In 2023, Sharma et al. compared feature detectors and descriptors on the suitability for image mosaicking, emphasizing the need for accurate keypoint detection so as not to add misalignments in the final mosaiced image. Their comparative study experimented with the efficiency of various combinations of detectors and descriptors based on the number of matched points, computation time, and overall image-stitching quality. Among all the combinations experimented with, the most versatile and efficient combination was identified to be the AKAZE-AKAZE combination in yielding high-quality stitched images with satisfactory computational performance [11].

In 2023, Lindenberger et al. presented Light Glue, a deep neural network that can recognize and compare local features in several photos. Examine a few of the design choices made by Superglue, the cutting edge of sparse matching, and extract small but powerful gains. Together, they improve Light Glue's accuracy, efficiency, and trainability in terms of memory and computing. First, Light Glue adapts to the difficulty of the challenge by making inferences considerably faster on image pairs that are intuitively easier to match, for example, because of high visual overlap or low appearance change. Thus, this creates extremely promising opportunities for the use of deep learning in latency-sensitive applications like 3D reconstruction [12].

3. INTEREST POINTS

An interesting point is the state of being recognizable and unique, at least locally. In other words, the point we find interesting will be easily distinguishable from its immediate surroundings. If we were to observe the image with a jeweler's glass, we would want to find an area we can easily recognize. Computationally, we achieve this by finding areas of multi-directional gradients in intensity around a single pixel in the image, characteristics most obviously seen in corners. A word about gradients. In an image, gradients are used to describe a change in intensity or color. A straight edge, like a wall, will show a significant gradient in a single direction, normal to the edge, denoting the change in intensity between the wall and its background in the image. Corners will exhibit gradients that radiate outward, starting normally from one of the constituent edges of the corner and sweeping around to the other. Clearly describe the interesting points that we have accumulated in each image. Since our interesting points are described as being locally distinguished via their gradients, we use the orientation of these gradients to help describe these points. We also look at the orientation of the gradients in a neighborhood around the points, extending our definition of "interesting" from part 1 (where we say a point is interesting if we can distinguish it within a small neighborhood) to include the neighborhood itself (i.e. using gradients to distinguish characteristics within small neighborhoods from other similar sized neighborhoods within the image). The algorithm to accomplish this consisted of convolving our image with directional filters to derive the gradients in each of 8 directions, then examining a patch around each interest point, weighted with a Gaussian centered on the interest point, to accumulate a histogram of gradients in each direction, in each of 16 bins around the interest point, for a total of 128 values in the descriptor for each interest point [13].

4. PROCESSING STAGE

This section includes a set of topics, the first of which is the data set, the Harris corner and its algorithm, as well as the SIFT descriptor and its algorithm, and finally, the matching features and algorithm that were explained.

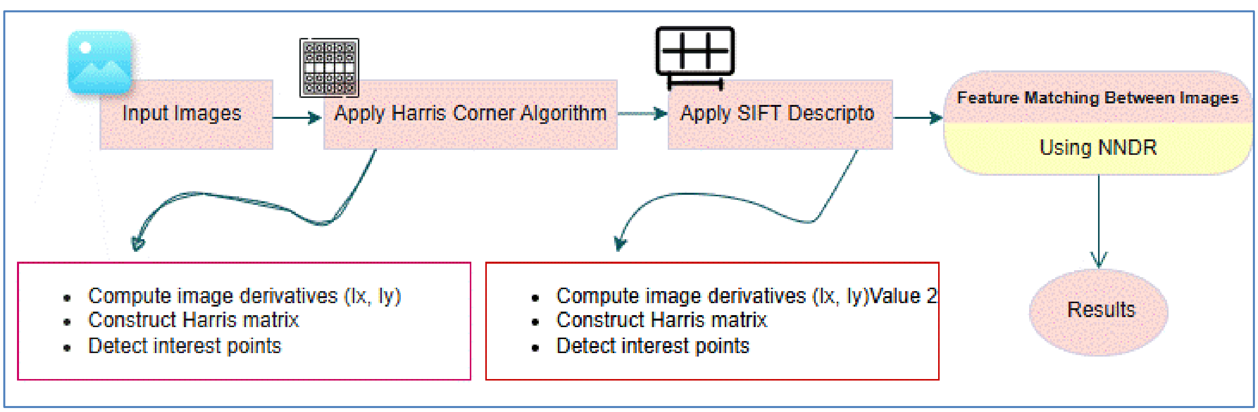


Figure 1: Processing Stage

4.1 Dataset

A data set of three of the most popular ancient monuments was employed in this research work to evaluate the performance of the proposed new feature detection and matching techniques. The selected monuments are the Ishtar Gate in Babylon, the Egyptian Pyramids in Cairo, and the Save Iraqi Cultural Monument in Baghdad. Two images of each monument were captured at two time instants to establish natural variations in illumination, viewpoint, and environmental conditions. This approach is supposed to emulate real-life situations in image matching. The dataset is structured to represent difficulty levels: the Ishtar Gate images are easiest due to well-defined structure features; the Egyptian Pyramids images are of moderate difficulty level; the Save Iraqi Cultural Monument images are the hardest due to challenging textures and less unique keypoints.

Figure 2: Ishtar Gate

Figure 3: Egyptian Pyramids

Figure 4: Save Iraqi Culture Monument

4.2 Harris Corner

Often employed in computer vision algorithms to extract corners and infer features from images, the Harris corner detector is a corner detection operator. Following the advancement of Moravec's corner detector in 1988, Chris Harris and Mike Stephens debuted the device. In contrast to using shifting patches at 45-degree angles, Harris' corner detector is more precise in differentiating between edges and corners, as it directly considers the differential of the corner score for direction. Ever since, numerous algorithms have included and enhanced it to prepare photos for later use. Discerning corners and edges. Ever since, numerous preprocessing methods have included and enhanced it [14].

One of the most often used corner detectors for extracting corners and identifying features from images is the Harris corner detector, which is combined with other computer vision techniques. Chris Harris and Mike Stephens first proposed the idea in 1988, building on Moravec's corner detector. Harris's corner detector proved to be more accurate in differentiating between edges and corners than the previous one because it takes the differential of the corner score for direction into consideration directly, rather than using shifting patches for each 45-degree angle. It has since been refined and used in numerous algorithms to prepare images for use in further stages of processing [14].

To remove the main elements of the picture, the Harris corner calculation is utilized. This calculation identifies the corners - interest utilizing the accompanying methodology:

4.2.1 Methodological steps of the Harris corner algorithm

1. **Gradient Estimation:** The first step involves computing the intensity gradients of the input image in both the horizontal (x) and vertical (y) directions, denoted as I_x and I_y, respectively. This is accomplished using Sobel operators in combination with a Gaussian smoothing filter to reduce the effect of noise and minor variations.
2. **Calculation of Gradient Products:** For every pixel in the image, the squared gradients I_x^2, I_y^2 , and the product of gradients I_x, I_y is computed. These operations are performed in a vectorized manner to enhance computational efficiency and minimize processing time.
3. **Gaussian Filtering of Gradient Products:** The gradient products are then smoothed by applying a Gaussian filter, resulting in S_{x^2} , S_{y^2} , and S_{xy}. This step helps aggregate information within a local neighborhood and is crucial for constructing the structure tensor at each pixel.
4. **Construction of the Harris Matrix:** Using the smoothed gradient values, the Harris matrix (also known as the second-moment matrix) is constructed at each pixel location. This matrix captures the intensity variation in the local neighborhood and is computed in a vectorized fashion for computational efficiency.
5. **Corner Response Function Computation:** The response of the Harris detector is calculated for each pixel using the following equation:

$$R = det\,(H) - k\,(trace\,(H))^2$$

Where *det (H)* and *trace (H)* are the determinant and trace of the Harris matrix, respectively, and (k) is a sensitivity factor typically set to 0.04 based on empirical evidence from prior studies.

6. **Non-Maximum Suppression and Interest Point Selection:** To identify distinct corners, non-maximum suppression is applied to the response map:
   - For each pixel, the corner response (R) is compared with a predefined threshold.
   - A 3×3 neighborhood around the pixel is evaluated.
   - If the pixel's (R) value exceeds the threshold and is the local maximum within its neighborhood, it is retained as a valid interest point.
   - The coordinates and corresponding gradient vectors of the selected points are stored for subsequent processing.

4.3 SIFT DESCRIPTOR

SIFT, Scale Invariant Feature Transform, is a computer vision algorithm for detecting and describing features. This finds those key points or features in images that become invariant to scale, rotation, and affine transformation. It functions by using local intensity extrema to identify key points and computing descriptors that capture the local image information surrounding those key locations. These descriptors can then be utilized in matching images, object recognition, and image retrieval tasks. A SIFT descriptor is a 3-D spatial histogram of the image gradient used to represent the appearance of the keypoint. Additionally, during model training, we can utilize the important points produced by SIFT as features for the image. The main benefit of HOG, over-edge, or SIFT features is that their effectiveness is independent of the image's size or orientation. The gradient at each pixel is like a small set of basic features of a vector, i.e., before carrying out further action by the operator, those features need to be formed on a three-dimensional (3D) coordinate and the angle of the edge. The weighted samples are used to compute the gradient norm according to the direction and induced in a 3-D histogram h, which (after normalization and clamping) will represent the SIFT descriptor. Another Gaussian function is further employed to diminish the contribution of more distant gradients from the central point [15]. The orientation encompasses eight bins, and the spatial coordinates encompass the first four, as follows:

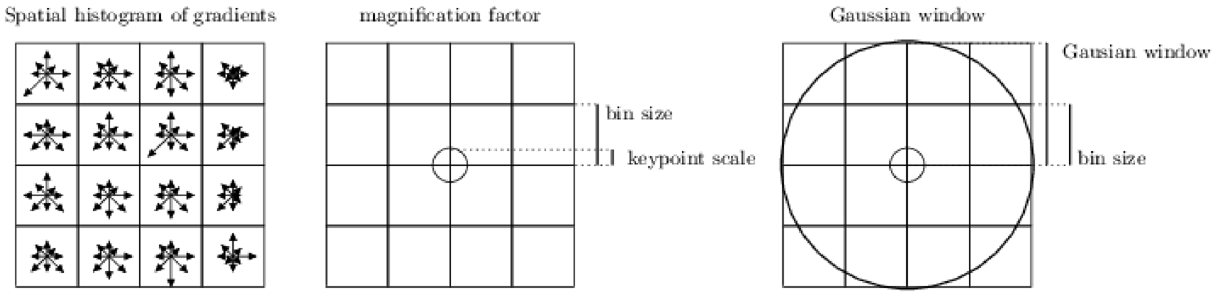


Figure 5: SIFT descriptor

4.3.1 Algorithm of SIFT Descriptor

The Scale-Invariant Feature Transform (SIFT) descriptor is constructed by following a sequence of computational steps to ensure invariance to scale, orientation, and illumination changes. The algorithm proceeds as follows:

1. **Orientation Estimation:** Compute the image gradient orientations using a combination of Sobel filters and Gaussian smoothing to suppress noise and stabilize gradient calculation.
2. **Dominant Orientation Assignment:** For each detected keypoint:
    a. A local neighborhood window of 256 pixels is extracted around the keypoint.
    b. A histogram of 36 orientation bins is computed within this region using quantization.
    c. The dominant orientation—i.e., the most frequent orientation value—is identified and subtracted from the keypoint's original orientation to ensure rotation invariance.
3. **Direction Quantization:** The orientation space is quantized into 8 principal directions to reduce complexity while preserving discriminative power.
4. **Feature Vector Construction:** The descriptor vector is built by analyzing the gradients within a 16×16-pixel window centered on the keypoint. This process involves:
    a. Initializing the start position 7 pixels before the keypoint in both row and column directions.
    b. Iteratively segmenting the window into 4×4 sub-regions (totaling 16 sub-regions).
    c. For each iteration:
        i. If the iteration index is divisible by 4, the row window is incremented by 4 pixels, and the column position is reset.
        ii. Otherwise, only the column position is incremented by 4 pixels.
    d. For each sub-region, compute an 8-bin histogram of gradient orientations based on the quantized direction values.
    e. These 8-bin histograms from each of the 16 sub-regions are concatenated, resulting in a 128-dimensional feature vector per keypoint.
5. **Normalization and Illumination Handling:**
    a. The resulting feature vector is first normalized to unit length.
    b. Values exceeding 0.2 are clipped to suppress the effect of large gradient magnitudes caused by illumination variations.
    c. The vector is normalized again to maintain robustness against lighting changes.
6. **Repeat for All Keypoints:**
   Steps 4 and 5 are repeated for each keypoint detected in the image to generate a complete set of SIFT descriptors for subsequent matching tasks.

4.4 Features Matching

The final step is to match the features. For each feature of image1, I calculated the distances between it and any feature from image2 and sorted the results. The distance represents the similarity of the features, and the ratio of the smallest distance and the second smallest one represents the reliability. If the ratio is greater than the threshold, I picked the pair as a match [16].

4.4.1 Algorithm of Features Matching

1. The number of features is the minimum number of features out of the two feature matrices.
2. Use MATLAB built-in function (bonus) to get the nearest two neighbors for each interest point in the first feature matrix from the second feature matrix.
3. Compute the nearest neighbor distance ratio (NNDR) by using an element-wise division between the first nearest neighbor vector and the second nearest neighbor vector.
4. Loop over the number of features and do the following:
   A. If the current value of the NNDR vector is less than a specific threshold (0.8):
      a) Add this value to the good match's vector.
      b) Add the current index of the feature1 matrix to the match's first column.
      c) Add the current index of the nearest neighbor matrix to the second column.
5. Sort the good matches so that the most confident matches are used in calculating the accuracy.

5. RESULT AND DECISION

This section presents the outcomes of applying the Harris Corner Detector and SIFT Descriptor, followed by the feature matching process. The analysis includes both quantitative results and qualitative discussions for each method and dataset.

5.1 Harris Corner Detection Results

Figures 6 to 8 illustrate the interest points extracted using the Harris Corner Detector for the three datasets. The performance of the Harris Corner Detector is influenced by the threshold set for interest point detection. Increasing this threshold results in fewer detected interest points, as only points with higher corner response values are retained. While reducing the number of interest points can improve computational efficiency, it may also reduce the robustness of feature extraction. Empirical results suggest that detecting between 2,000 to 15,000 interest points provides a practical balance between accuracy and resource consumption. A higher number of key points generally leads to improved matching accuracy but increases memory usage and processing time.

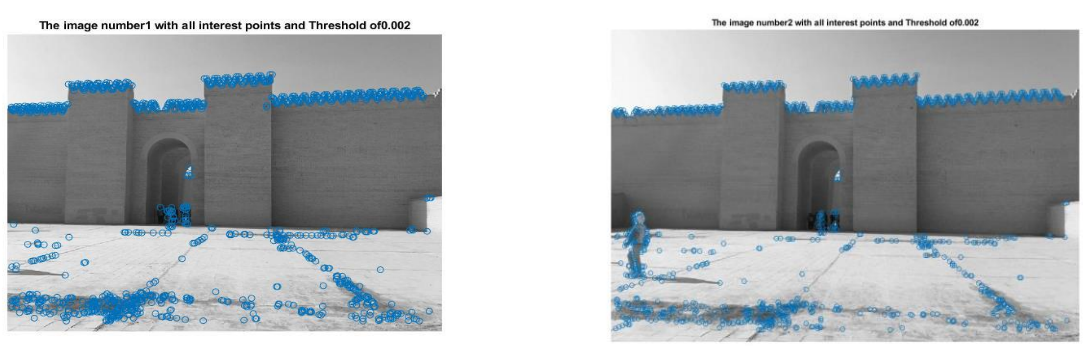


Figure 6: The interest point of Ishtar Gate.

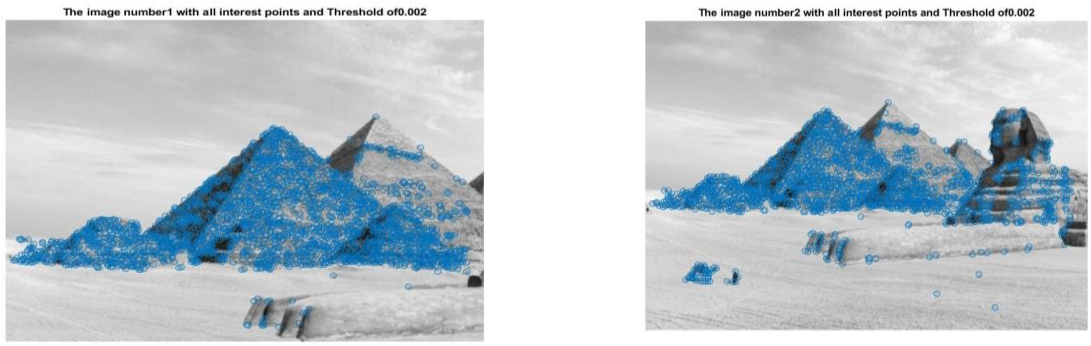


Figure 7: The interest point of Egyptian pyramids.

Figure 8: The interesting point of the Save Iraqi Culture monument

5.2 The result of the SIFT Descriptor

The SIFT descriptor generates a matrix of size *k×n*, where k is the number of keypoints and n is the 128-dimensional feature vector for each keypoint. This transformation enables scale and rotation-invariant representations for each interest point.

SIFT descriptors are inherently invariant to rotation due to the assignment of a dominant orientation to each keypoint. Furthermore, the descriptor is robust to changes in illumination, as it includes normalization and thresholding mechanisms that minimize the impact of varying light conditions. These properties make SIFT suitable for diverse real-world applications involving image matching and object recognition.

5.3 The Result of Features Matching

Figures 9 to 11 display the feature matching outcomes for each of the three monuments using the combined Harris and SIFT pipeline:

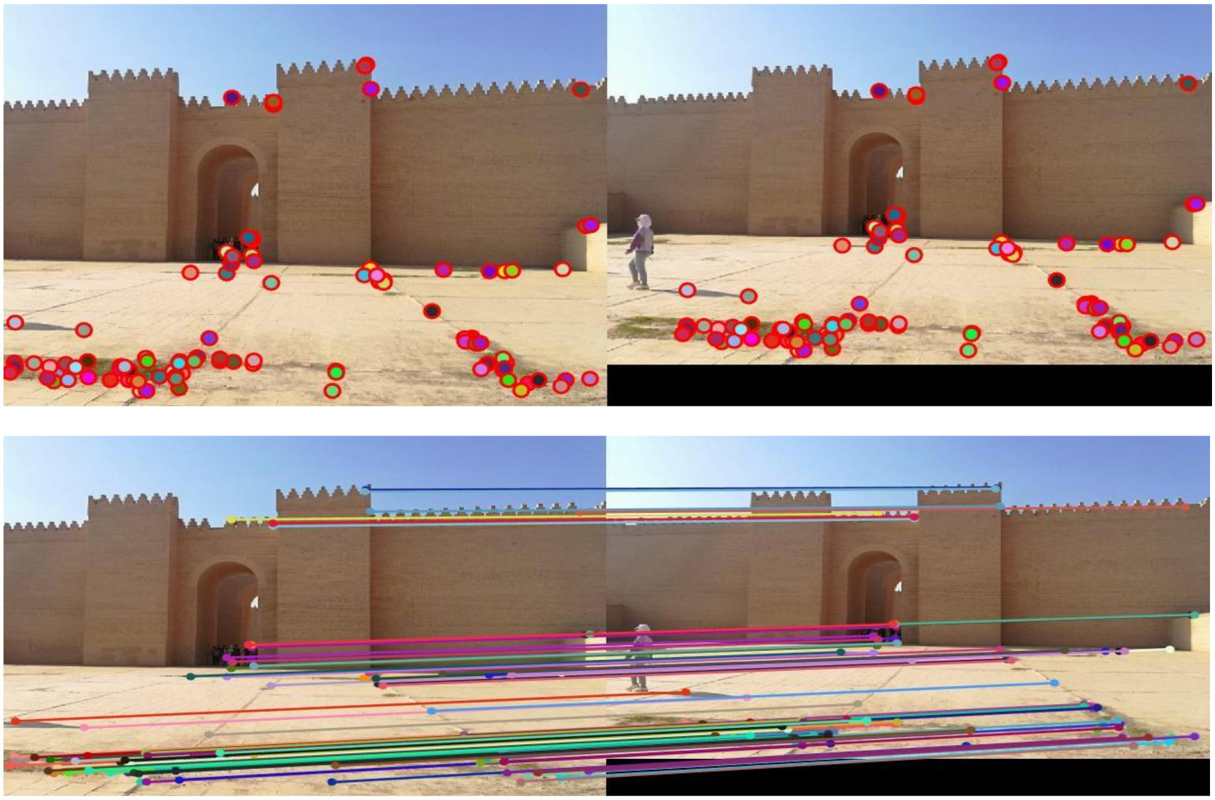
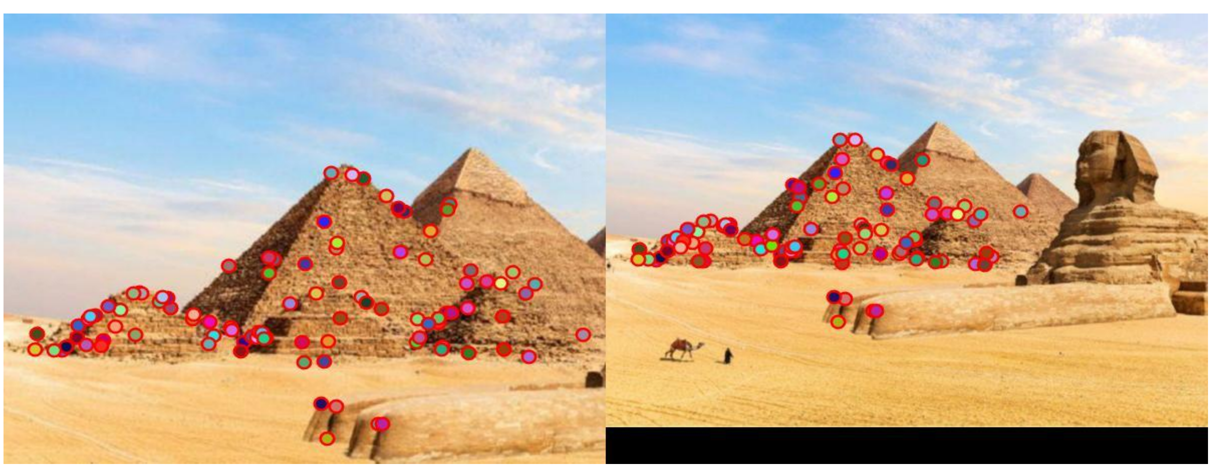


Figure 9: Features Matching of Ishtar Gate



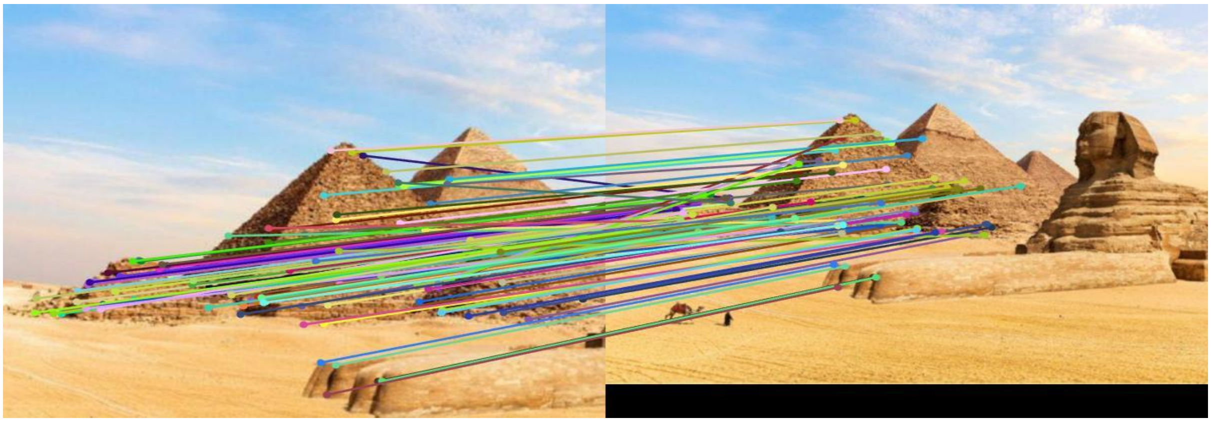


Figure 10: The Features Matching of Egyptian pyramids.

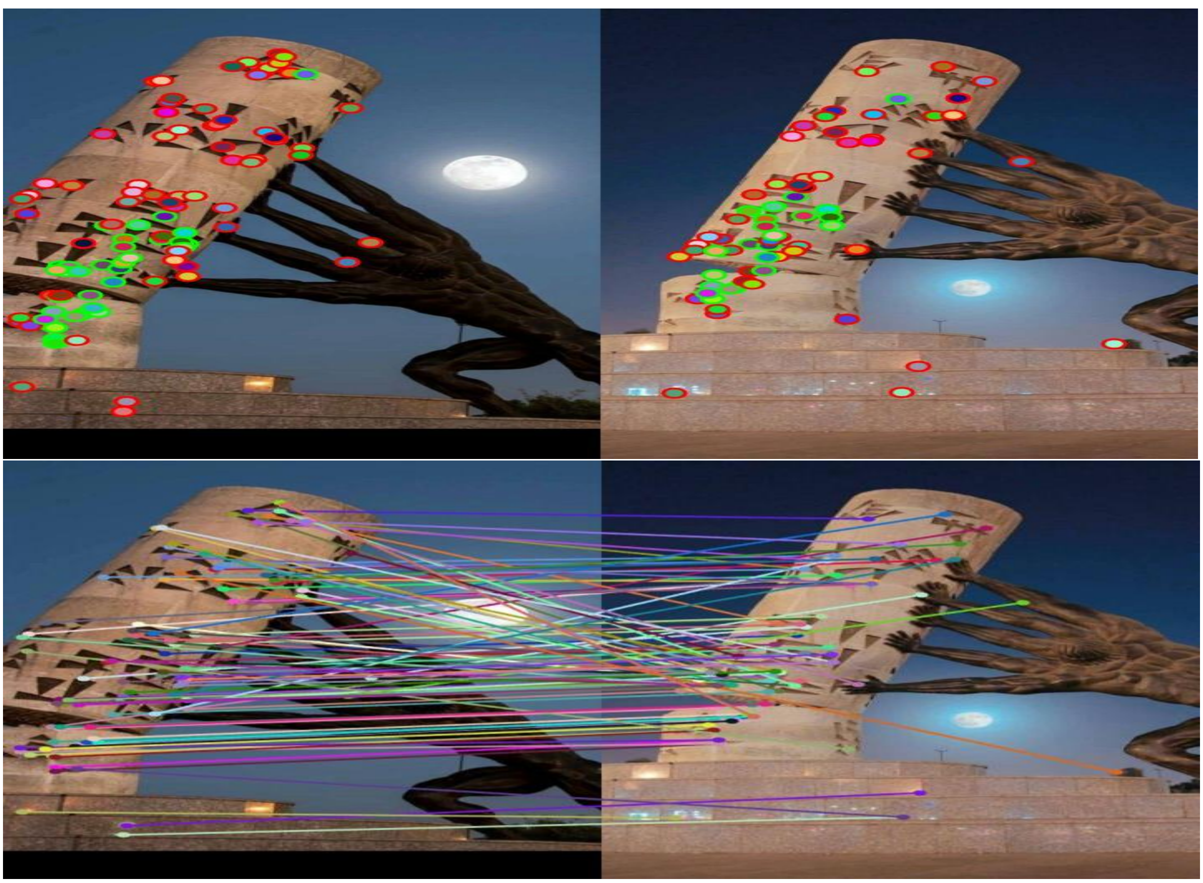


Figure 11: The Features Matching of the Save Iraqi Culture Monument

Table 1: Feature Matching Results for the Three Datasets

Figure	Dataset	Good Matches	Bad Matches	Precision (%)	Accuracy (Top 100) (%)
Figure 8	Ishtar Gate	99	0	100.00	99.00
Figure 9	Egyptian Pyramids	99	1	99.00	99.00
Figure 10	Save Iraqi Culture Monument	94	6	95.00	95.00

The results in Table 1 demonstrate the high effectiveness of the Harris-SIFT feature matching pipeline. Both the Ishtar Gate and Egyptian Pyramids datasets achieved near-perfect precision and accuracy, indicating strong geometric consistency and lighting conditions between the image pairs. However, the Save Iraqi Culture Monument dataset yielded slightly lower precision and accuracy, likely due to scale and texture variations between the images. These findings underscore the importance of scale-invariant implementations in robust image matching systems.

6. CONCLUSION AND FUTURE WORK

This research integrates the Harris corner detector with the SIFT descriptor to achieve strong local feature matching across images. The findings indicate that high accuracy in matching is realized, particularly in organized images, with the validity of this approach established across numerous imaging conditions. The combination of Harris and SIFT proves very effective in video surveillance scenarios, where accurate tracking of objects or events across successive frames is required. Besides, the enhanced real-time performance—achieved by reducing false matches and maximizing feature extraction—allows its usage in smart systems and real-time scenarios, such as autonomous vehicles or surveillance security. Its handling in challenging scenes and real-time video streams can be improved in future work by parallelizing computations or integrating deep learning with the Harris-SIFT pipeline.

Reference

[1] Fu, Y., Lei, Y., Wang, T., Curran, W. J., Liu, T., & Yang, X. (2020). Deep learning in medical image registration: a review. Physics in Medicine & Biology, 65(20), 20TR01.

[2] Bekkanti, A., Karimunnisa, S., Gogulamudi, S., Kumar, K. T. P. S., & Basha, C. Z. (2020, September). Enhanced computerized bone fracture detection using Harris corner detection. In 2020 International Conference on Smart Electronics and Communication (ICOSEC) (pp. 572-576). IEEE.

[3] Harris, C. G., Jursik, N. J., Rochefort, W. E., & Walker, T. W. (2019). Additive manufacturing with soft TPU–adhesion strength in multimaterial flexible joints. Frontiers in Mechanical Engineering, 5, 37.

[4] Bojanić, D., Bartol, K., Pribanić, T., Petković, T., Donoso, Y. D., & Mas, J. S. (2019, September). On the comparison of classic and deep keypoint detector and descriptor methods. In 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA) (pp. 64-69). IEEE.

[5] J. Qin et al., "An Encrypted Image Retrieval Method Based on Harris Corner Optimization and LSH in Cloud Computing," IEEE Access, vol. 7, pp. 24626–24633, 2019, doi: 10.1109/ACCESS.2019.2894673.

[6] Feng, J., Ai, C., An, Z., Zhou, Z., & Shi, Y. (2019, July). A feature detection and matching algorithm based on the Harris Algorithm. In 2019 International Conference on Communications, Information Systems and Computer Engineering (CISCE) (pp. 616-621). IEEE.

[7] Efe, U., Ince, K. G., & Alatan, A. (2021). Dfm: A performance baseline for deep feature matching. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4284-4293).

[8] F. Bellavia and D. Mishkin, "Harris Corner Selection for Next-Gen Image Matching Pipelines," no. 1981, pp. 1–8, 2021, [Online]. Available: http://arxiv.org/abs/2109.12925

[9] F. Bellavia and D. Mishkin, "HarrisZ$^+$: Harris Corner Selection for Next-Gen Image Matching Pipelines," no. 1981, pp. 1–8, 2021, [Online]. Available: http://arxiv.org/abs/2109.12925.

[10] Wang, Z., Yu, A., Dong, Z., Zhang, B., & Chen, X. (2022). Performance Evaluation of Interest Point Detectors for Heterologous Image Matching. Remote Sensing, 14(15), 3724.

[11] Sharma, S. K., Jain, K., & Shukla, A. K. (2023). A comparative analysis of feature detectors and descriptors for image stitching. Applied Sciences, 13(10), 6015.

[12] Lindenberger, P., Sarlin, P. E., & Pollefeys, M. (2023). Lightglue: Local feature matching at light speed. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 17627-17638).

[13] He, R., Chen, H., & Fu, M. (2024). Interest Point Selection and Feature Extraction in Six-DoF Grasp Pose Detection. IEEE/ASME Transactions on Mechatronics, 29(6), 4723-4733.

[14] Derpanis, K. G. (2004). "The Harris Corner Detector." York University, Department of Computer Science and Engineering.

[15] Y. Chen, S. Shen, Y. Chen, and G. Wang, "Graph-based parallel large scale structure from motion," Pattern Recognit., vol. 107, no. xxxx, 2020, doi: 10.1016/j.patcog.2020.107537.

[16] Xue, F., Zhou, Z., Ye, W., Guo, Y., & Wang, Y. (2025). MATCHA: Towards Matching Anything. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2025), 12345–12355. IEEE DOI: https://doi.org/10.48550/arXiv.2501.14945.