Amalgamation of ORB and SWIFT to identify an image efficiently

[1*]Manish Kumar Thakur and [2]Dr. Raju Ramakrishna Gondkar

[1]Research Scholar, Visvesvaraya Technological University, Belgaum*
[2]Research Guide, Visvesvaraya Technological University, Belgaum

Email Address: [1]manishkumarthakur2023@gmail.com, [2]rajuramakrishnagondkar@gmail.com*

Abstract: It seen in the recent times that huge data of images is common due to exposure of digital camera to society. This creates a requirement of processing this huge data repository based on a few common requisites. Number of algorithms is for various different of processing seen in use. This paper provides the comparison of ORIENTED fast and ROTATED BRIEF (ORB) and SCALE INVARIANT FEATURE TRANSFORM (SIFT) methods for different scaled and rotated values. A study can 0064i date image with different values for rotation and scaling considered finding the most efficient combination to be use. The paper also provides a combination of ORB and SIFTS method, which provides the most efficient results. The analysis of results yields novel results indicating the optimization of computation time and accuracy of identification. The paper brings about an identification method where in an image can be rotated to certain degree based on the candidate image position and then identified to provides more time efficient identification.

Keywords: ORB, SWIFT, Rotation Precision, Scaling Precisions, Machine Time

1. INTRODUCTION

In the last decade, large database of images have grown rapidly and allowed machine to understand different ways of human vision. Image retrieval and query to understand the different visual aspects of these large databases are the most important aspects of the computer vision. In order to find an image, the different image feature representation like color, texture, orientation, and size are the most important aspects. Human eyes are able to differentiate these aspects in much better way without difficulties. However, to make machine understand extraordinary capabilities for Computer Vision applications like AR/VR, medical images, satellite image, etc. needs improvise image retrieval and hybrid algorithms.

We are adding huge number of digital images every second in the world. We are also now willing to be in the virtual world with augmented and virtual reality. The new need with the different scaled and rotated images for AR & VR have pushed for hybrid and improved feature detection and description technology. The quick changing imaging condition, surface, viewpoint, scale, brightness, accuracy, resolution, energy consumption, storage, and the rendering devices have allowed the birth more hybrid approach with a combination of ORB and SIFT.

An ordinary human skill for object recognition presents an exceptional challenge for computer vision systems. We are trying to analyses the different capabilities and challenges for the ORB and SIFT. The paper concludes the findings for ORB and SIFTS which will allow us to develop a hybrid image retrieval system for the future.

The rapid development of development of digital image processing units with growing capabilities have provided a challenge to upgrade the existing

Techniques to suit for the growing needs of fast and accurate feature extraction. The development of AR/ VR has created a need for modification of existing techniques so; it may be used to the changed requirements. In a paper we are trying to implement, the existing ORB and SIFT with different orientation and shape of the image. ORB and SIFT are one of the most suitable feature extraction methods for the current needs of computer vision. However, both of the methods are having their props and cons. We are trying to find a combination of both which can be able to fulfil the needs of current computer vision. We have analysed the effectiveness and accuracy of ORB and SIFT for the feature detection, extraction, and matching with different orientation and shape. We have also compared the performance of both approaches and presented a way for the development of hybrid image retrieval system.

2. ORIENTED FAST AND ROTATED BRIEF (ORB)

The need for accurate and alternative method for SIFT and SURF was felt, which lead to ORB. ORB is a combination of Features from Accelerated Segment Test (FAST) and Binary Robust Independent Elementary (BRIEF). FAST technique used to locate key point and BRIEF for descriptor extraction process. When ORB implemented both FAST abd SWIFT are considered. Rotation invariance achieved by ORB using feature detection algorithm applied to intensity centroid method to FAST. ORB is twice as fast as SIFT and much faster than SURF. Orb has improved over the years and this improvement largely attributed to optimization of original real-time calculation.

2.1 Using direction information for Feature Points Detection

FAST applied to the candidate image repeatedly to layers of the pyramid to get scale invariance. Excluding the non-corner features, Harris measurement used as testing method to check the correctness of detection of FAST features. Lack of directional information and the high computational advantages are the main reasons for use of FAST algorithm. ORB uses effective grey centroid method to describe the features point's direction which id hence called oFAST. FAST as the name indicates is quick in computation and is a corner point that is useful I n identifying local pixels which have significant grey-scale changes. Hence, FAST helps in comparing the grey values of pixels, which leads to other problems like the lack of direction information. ORB overcomes this disadvantage by using geometric moment of the image to provide feature points with direction information to get grey centroid in the neighborhood of the given radius. The main direction of the feature point is the vector from the feature point to the grey centroid. However, FAST does not suffer from rotation invariance.

2.2 Feature Points Descriptors with rotation angle

Considering a group of n pixels as a patch, an n-bit string produced by BRIEF algorithm. The instability of this vector against rotation provides an opportunity for improvement. The issue addressed by ORB and uses rotation aware variant of BRIEF. ORB provides features in a rotation-invariant form. ORB for the located key points, first defines a patch centroid based on the image moments. The direction of the vector from key point centre to its patch centroid provides the direction. This binary test pattern can rotated to align the direction of the patch before extracting the descriptors. This allows a feature to independent of rotation.

ORB algorithm gets its descriptor by rBRIEF (rotated BRIEF) algorithm, which an algorithm working based on the above description, which is independent of rotation. Brief selects a specific number of pixel point pairs surrounding the feature point of interest. The algorithm compares the grey values of these pixel pairs and produces a string of binary features descriptors. rBRIEF provides the improvement need for BRIEF to be rotation independent.

2.3 Feature Points Matching

ORB uses the Hamming distance to compare the similarities between the feature points. The similarity measurement calculated using the XOR of the binary feature string. The efficiency of this similarity measurement is very high.

3. Scale Invariant Feature Transform (SIFT)

This is also a feature detection algorithm, which used to extract local features of the image and identify the image. It is a key point detection and descriptor extraction technique. This algorithm does not vary based on scaling, brightness, rotation, noise, affine transformations or any other perspective transformation. The algorithm uses set of reference points and extracts key points of candidate image. These key points are stored in a database. This algorithm identifies the image from a given image by comparing each feature of the new image with the features stored in the database and then locating the candidate-resembling image. SWIFT uses Euclidean distance for matching the feature vectors. The algorithm provides a scale invariant feature vector by constructing a pyramid of the input image and by applying Gaussian operation to identify the local extrema in the scale space. This results in a 3 X 3 X 3 window. The next step is to identify the key points that are not in the extrema but have high contrast. The image gradients of a 16 × 16 window, centered at each key point, then computed and grouped into 4×4 sub regions. The direction of the gradient then computed for each sub region. This represented by an eight-bin histogram. The summation of all the bins from sixteen histograms in the window is done and this 128 element SIFT descriptor vector is inferred.

3.1 Existing work

A number of feature descriptive algorithms like the Local Binary Patter, SURF, oFAST, ORB used to solve the classification problem in an image. Number of algorithms to calculate the accuracy of theses classifications also used. It is inferred that feature descriptor algorithms work more powerful on original images rather than Histogram Equalized images and the CNN based proposed model that is trained on Histogram Equalized image dataset gives the best accuracy of 97.32 %. [13]

A method to solve the problem of low accuracy of feature matching in oFAST and rORB, a mismatching elimination algorithm designed. A preliminary testing using cosine method for a pairs of descriptors identified by Brute force method. The method used one more algorithm called the Random Sample Consensus (RANSAC), which works on homography matrix and removed further mismatches. Experiments have proved this is an efficient method and is successful in removing mismatching points. This provided better accuracy in matching when compared to original ORB. Applications with real-time performance and that need high accurate matching are best fit for this method.

4. ORB Vs SIFT Comparison Experimental Steps And Values

4.1 Pre-processing steps:

1. Rotate and Scale the image to provide different types of inputs nearer to the real world to the experiment. In addition, we are

trying to find the difference in efficiency of ORB and SIFT differs during the rotation and scaling of image.

2. Rotation of image to compare - The image to be compared was rotated from 0 to 300 in steps of60-degree angles with bounding (*60 degree considered because of no significant change inmatches within 60-degree changes)
3. Scaling of image to compare - The Image to compare as scaled from 30% of source image to180% of the same in steps of 30% changes (*30% scaling considered because of no significant change in matches within 30% scaling)

4.2 Tools and Technologies

1. Python
2. OpenCV Libraries
3. Supporting libraries for OpenCV, ORB, SWIFT
4. Hardware
   a) Basic system with i3 and above having minimum graphics card,
   b) 5 GB to store the programs and other supporting files,
   c) Minimum of 8GB RAM for smooth run.

5. EXPERIMENTAL DESIGN

5.1 Tests and values for ORB

a) Detecting the key points and descriptors from the source image and note, the time required.
b) Detecting the key points and descriptors of pre-processed images (scaled and rotated images) and note the time required.
c) Match the features from source image and image to compare and note the time required for matching.
d) The matches returned contain both correct matches and wrong matches, so the matches to curate based on Distance between descriptors matched in brute force match using NORM_HAMMING.
e) During the experimental approach different values, for the threshold distances were tested beyond which the number of wrong matches significantly increased and hence dropped, so a distance of 40 was experimentally determined as the optimal threshold to curate the matches.
f) The time taken to curate the matches also documented.
g) The total time taken for the orb to finish detecting, matching and curating the matches also considered for further comparison.

5.2 Tests and values for SIFT

h) Detecting the key points and descriptors from the source image and note the time required
i) Detecting the key points and descriptors of pre-processed images (scaled and rotated images) and note the time required.
j) Match the features from source image and image to compare and note the time required for matching.
k) The matches returned contain both correct matches and wrong matches, so the matches to curated based on the ratio test distances returned by Brute forceknn Match.
l) During the experimental approach different ratios of distances were tested below which the number of wrong matches significantly increased and hence dropped, so a distance ratio of 0.75 was experimentally determined as the optimal threshold to curate the matches.
m) The time taken to curate the matches also documented.
n) The total time taken for the orb to finish detecting, matching, and curating the matches is also considered for further comparison

5.3 Additional steps

o) Calculating True positive: The true positives calculated by removing outlier points matched from the total matched features.
p) Calculating False positive: The count of outliers removed by threshold-based curation in both ORB and SIFT feature matching.
q) Calculating percentage of completion (Precision):It calculated as the number of true positives being correctly matched features to the combined sum of correctly matched features and the number of outliers removed.

Table-1: Comparison of ORB and SIFT [Boat-Rotation]

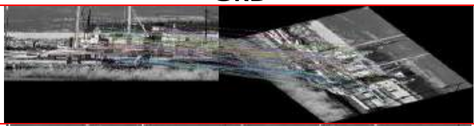
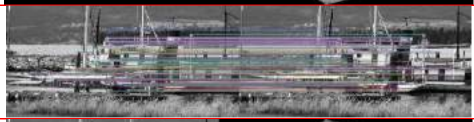
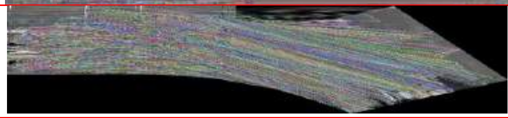
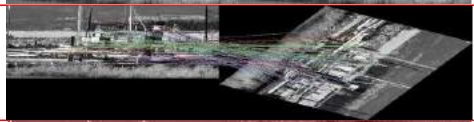
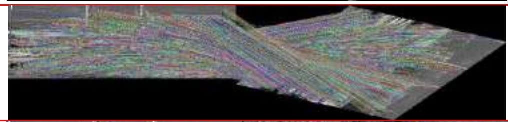
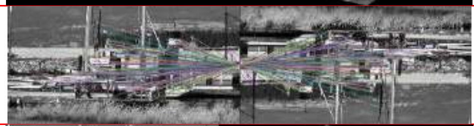
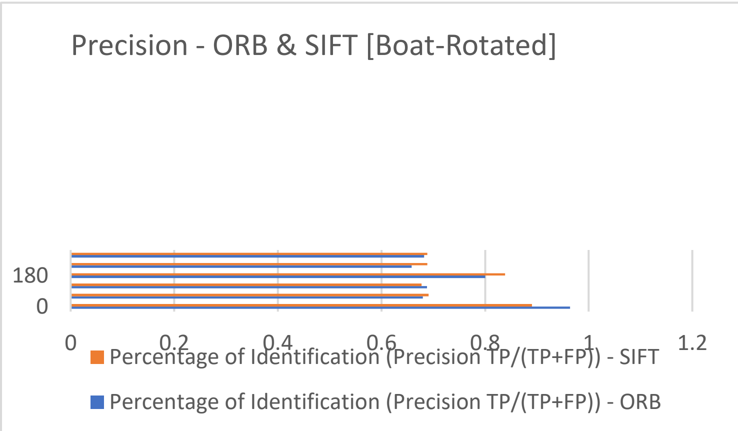
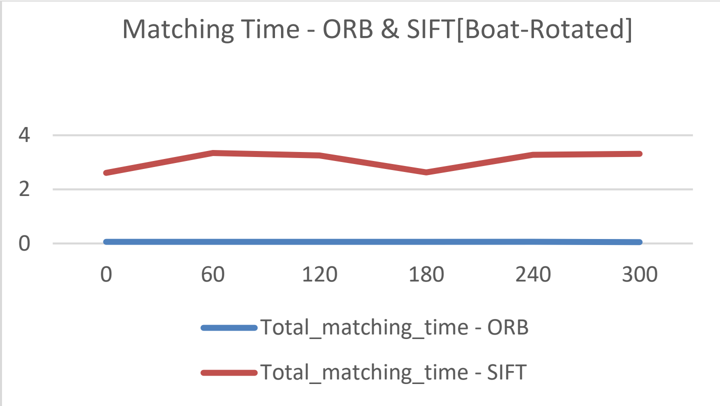
Comparison of ORB and SIFT (Boat– Rotation)	
ORB	SIFT
	
	
	
	
	
	

Table-2: Rotation with Precision and Matching Time

Rotated	Percentage of Identification (Precision TP/(TP+FP)) - ORB	Percentage of Identification (Precision TP/(TP+FP)) - SIFT	Total_matching_time - ORB	Total_matching_time - SIFT
0	0.964	0.89061445	0.062474966	2.609173059
60	0.68	0.690636957	0.06249404	3.343498468
120	0.688	0.77245105	0.062496424	3.249750614
180	0.8	0.838510016	0.062472343	2.624799728
240	0.658	0.688161152	0.062500954	3.281004429
300	0.682	0.688273689	0.049805641	3.308860302





Graph-1: Rotation Precision and Matching Time ORB & SIFT

Table-3: Comparison of ORB and SIFT [Boat Scaling]

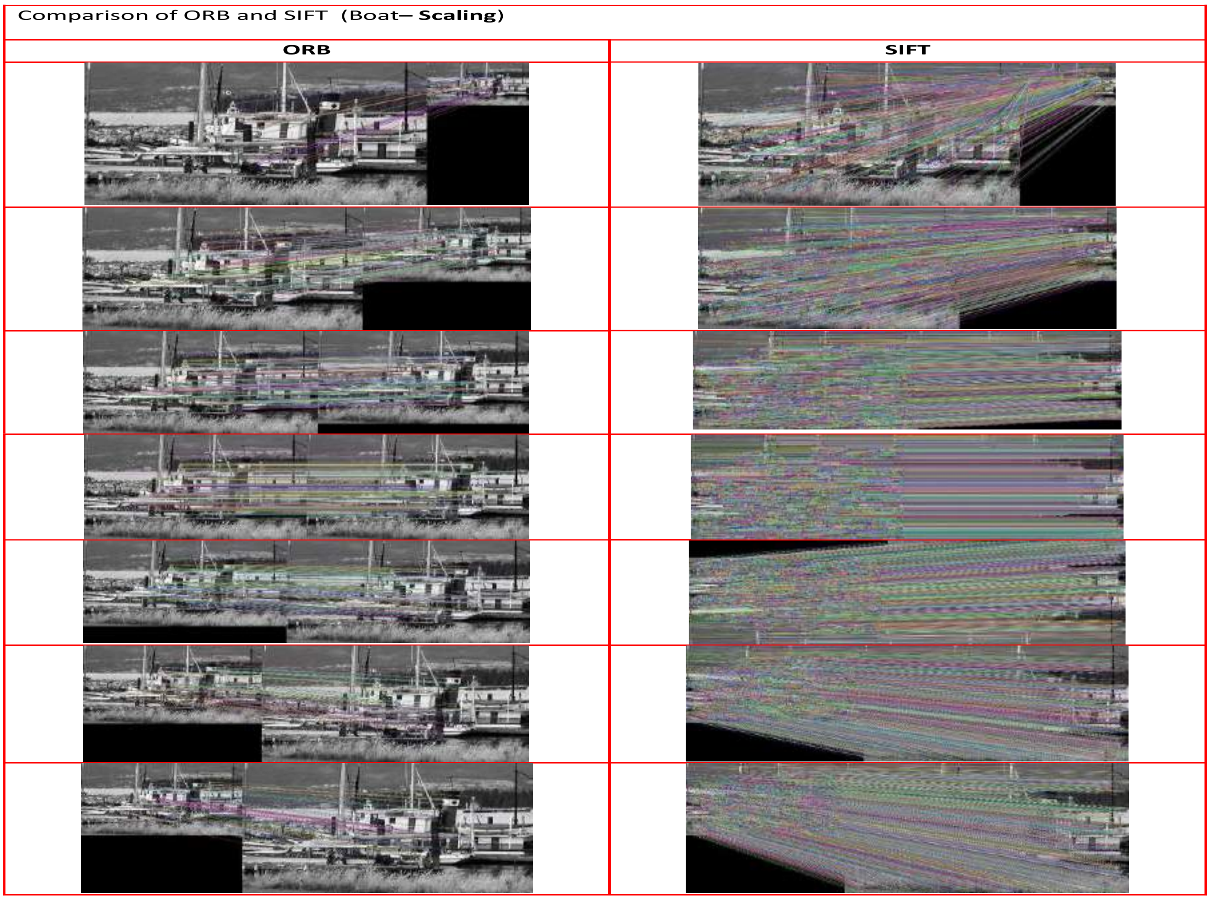
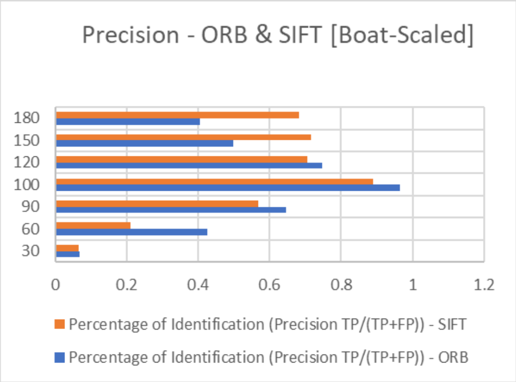
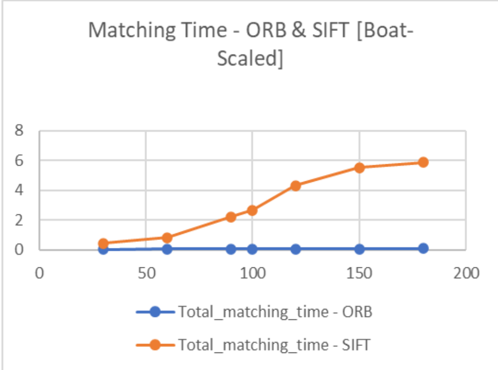


Table-4: Scaling with Precision and Matching Time

Scaled	Percentage of Identification (Precision TP/(TP+FP)) - ORB	Percentage of Identification (Precision TP/(TP+FP)) - SIFT	Total_matching_time - ORB	Total_matching_time - SIFT
30	0.068	0.066284042	0.031253099	0.453089714
60	0.426	0.210330858	0.046853781	0.828066111
90	0.646	0.568759847	0.04597187	2.216287613
100	0.964	0.89061445	0.046849728	2.671692133
120	0.746	0.704253883	0.062496901	4.335020065
150	0.498	0.715395003	0.062492132	5.51385355
180	0.406	0.681071348	0.093742132	5.867728949





Graph-2: Scaling Precision and Matching Time ORB & SIFT

Table-5:Comparison of ORB and SIFT [Villa-Rotation]

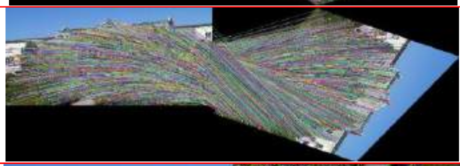
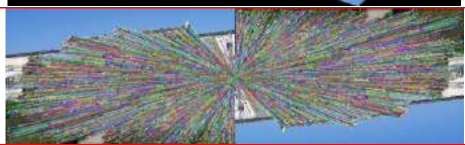
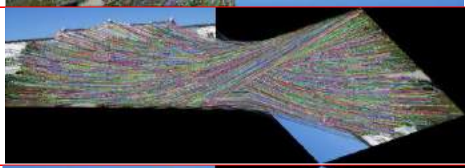
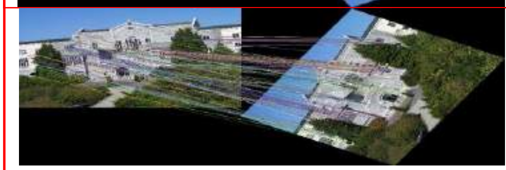
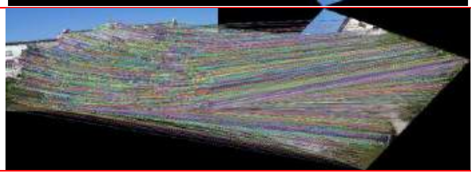
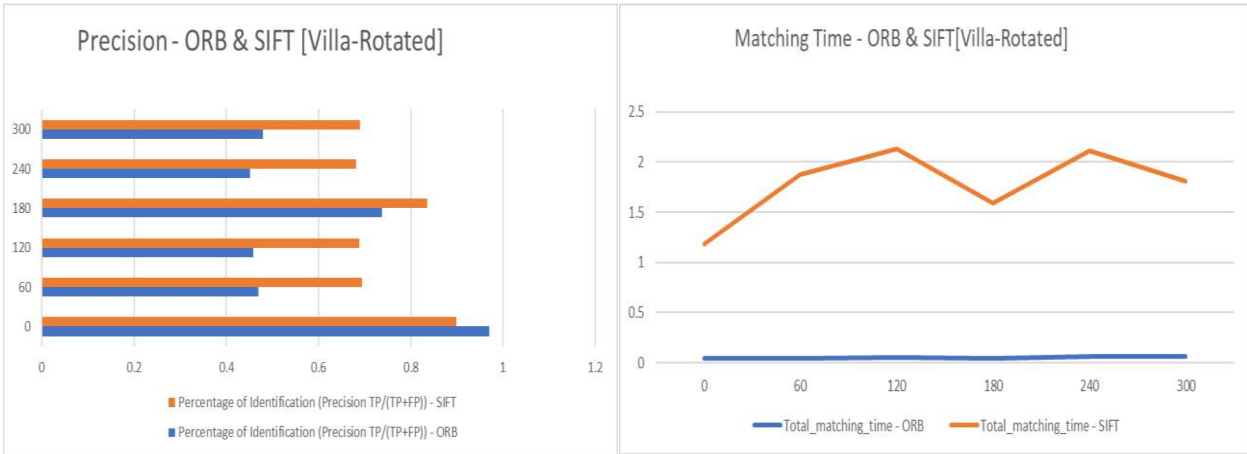
Comparison of ORB and SIFT (Villa— **Rotation**)	
ORB	**SIFT**
	
	
	
	
	
	

Table-6: Rotation with Precision and Matching Time

Rotated	**Percentage of Identification (Precision TP/(TP+FP)) - ORB**	**Percentage of Identification (Precision TP/(TP+FP)) - SIFT**	**Total_matching_time - ORB**	**Total_matching_time - SIFT**
0	0.97	0.898236776	0.045974016	1.179072142
60	0.47	0.694038623	0.046871901	1.871914387
120	0.458	0.688664987	0.050969601	2.136349916
180	0.738	0.835264484	0.046871901	1.592091799
240	0.452	0.681276238	0.060446501	2.111105204
300	0.48	0.690176322	0.062479258	1.804891109



Graph-3: Rotation Precision and Matching Time ORB & SIFT

Table-7: Comparison of ORB and SIFT [Boat Scaling]

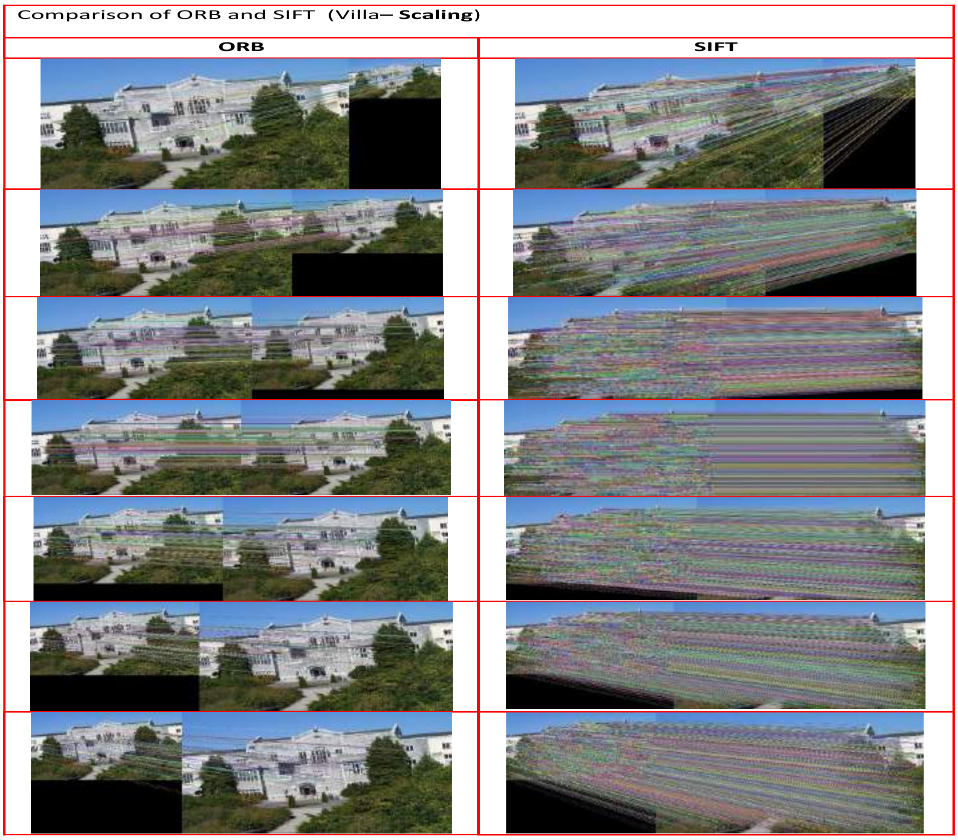
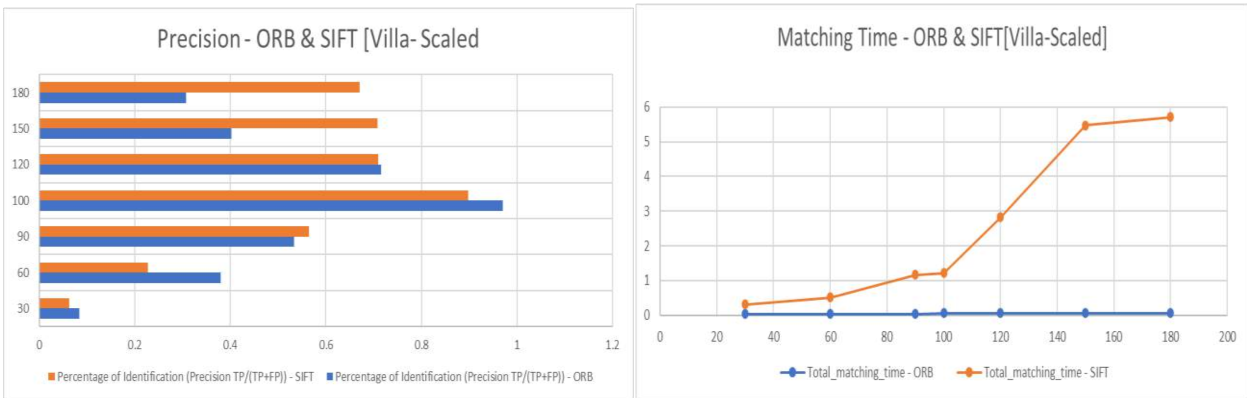


Table-8: Scaling with Precision and Matching Time

Scaled	Percentage of Identification (Precision TP/(TP+FP)) - ORB	Percentage of Identification (Precision TP/(TP+FP)) - SIFT	Total_matching_time - ORB	Total_matching_time - SIFT
30	0.084	0.062468514	0.031251669	0.312473059
60	0.38	0.227371956	0.031251669	0.515572548
90	0.534	0.564399664	0.038972616	1.170750856
100	0.97	0.898236776	0.062496662	1.218666315
120	0.716	0.709319899	0.062493324	2.826225519
150	0.402	0.708480269	0.062496185	5.474201679
180	0.308	0.670696893	0.062498093	5.707921505



Graph-4: Rotation Precision and Matching Time ORB & SIFT

5.4 Rotation Data Analysis-

1. Base image – 0 degree
2. Candidate image – 0 to 300 [gap 60]
3. 'rotated_angle_candidate_image'- Angle of rotation of candidate Image (ranging from 0 to 360)
4. 'Number_of_Features_matched' (TP)- Number of features matched with respect to base and candidate images
5. 'outliers_rejected_if_any' (FP)- Number of wrongly matched features rejected based on threshold (40 in orb and 0.75 in sift)
6. 'Percentage of Identification (Precision)-
   Precision = TP/ (TP+FP)
   Where,
   1. Precision = The percentage of identification
   2. TP = True positive represents the number of correct features matched "Positive"
   3. FP = False positive represents the outliers rejected if any "Negative"
   4. 'Total_matching_time'- Total time for matching and curating the correct matches.
   5. t5_fin = t1_fin+t2_fin+t3_fin+t4_fin

Where,

t5_fin = 'Total_matching_time'

t1_fin = Time taken to detect feature points in base image

t2_fin = Time taken to detect feature points in candidate image

t3_fin = Time taken to match the features with respect to base and candidate image

t4_fin = Time taken to eliminate wrongly matched features based on threshold

5.5 Scaled Data Analysis-

1. Base image – 100
2. Candidate image – 30 to 180 [gap 30]
3. 'rotated_angle_candidate_image'- Angle of rotation of candidate Image (ranging from 30 to 180)
4. 'Number_of_Features_matched' (TP)- Number of features matched with respect to base and candidate images
5. 'outliers_rejected_if_any' (FP)- Number of wrongly matched features rejected based on threshold (40 in orb and 0.75 in sift)
6. 'Percentage of Identification (Precision)-
   Precision = TP/ (TP+FP)
   Where,
   - Precision = The percentage of identification
   - TP = True positive represents the number of correct features matched "Positive"
   - FP = False positive represents the outliers rejected if any "Negative"
7. 'Total_matching_time'- Total time for matching and curating the correct matches.

t5_fin = t1_fin+t2_fin+t3_fin+t4_fin

Where,

t5_fin = 'Total_matching_time'

t1_fin = Time taken to detect feature points in base image

t2_fin = Time taken to detect feature points in candidate image

t3_fin = Time taken to match the features with respect to base and candidate image

t4_fin = Time taken to eliminate wrongly matched features based on threshold

5.6 Findings of Data Analysis

1. ORB is faster than SIFT in all inputs for candidate image.
2. ORB performs well during scaling of 60 to 120
3. SIFT performs well before 60 and above 120 scaling
4. ORB performs well on 0 and 180 degrees
5. SIFT performs well on 60, 120, 240, and 300 degree

6. CONCLUSION

The objective of the work was to compare the ORB and SIFT methods for different scaled and rotated values. The need to provide the candidate image with different values for rotation and scaling to find the most efficient combination to use was due to the need to understand the various parameters of identification. The paper develops a combination of ORB and SIFTS which provides the most efficient results. As per the results, it can be conclude that ORB is faster and SIFT is more accurate. Considering the facts that ORB is the fastest algorithm with a smaller number of feature points while SIFT performs the best with a greater number of feature points in the most scenarios with more time so, we have developed a hybrid model along with optimal parameter valueswith the two algorithms to have better computational efficiency and accuracy. The methods provides the best combination results for the rotated and scaled candidate images in the case of ORB and SIFT. The method determines the optimal values for rotation and scaling where the two algorithms are best suited for better image retrieval in the case of real-time applications. The method considers the two important preprocessing aspects for image retrieval i.e., scaling and rotation to find the best scenarios for both feature detection and matching algorithms. The work has clearly indicated that both the algorithms are having their advantages and disadvantages. The combination with correct preprocessing elements will provide better results.

ACKNOWLEDGMENT

Funding

Not Applicable

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Authors' contributions

Dr. Raju Ramakrishna Gondkar supervised every step of the work and provided critical review and valuable input. All authors read and approved the final manuscript.

REFERENCES

1. E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF." 2011 International Conference on Computer Vision, Barcelona, pp. 2564-2571, 2011.
2. E. Rosten and T. Drummond, "Machine learning for high speed corner detection." In Proc. European Conf. Computer Vision, vol. 1, pp. 430– 443, 2006.
3. M. Calonder, V. Lepetit, C. Strecha, and P. Fua," BRIEF: Binary robust independent elementary features." In Proc. European Conf. Computer Vision, pp. 778–792, 2010.
4. J. Gao, Y.F Wu, K. Wu. "Image Matching Algorithm Based on Corner Detection [J]."Chinese Journal of Scientific Instrument, vol. 34, no. 8, pp. 1717-1725, 2013.
5. M. Calonder, V. Lepetit, C. Strecha, et al. "BRIEF: binary robust independent elementary features[C]"European Conference on Computer Vision,pp. 778-792, 2010.
6. E. Rosten, R. Porter, T. Drummond."Faster and Better: A Machine Learning Approach to Corner Detection [J]."IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 32, NO. 1, PP. 105-119, 2010.
7. R. Mur-Arta, J. M. M. Montiel,J. D. Tardos J D. "ORB-SLAM: a versatile and accurate monocular SLAM system[J]."IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, 2015.
8. C. Harris and M. Stephens, "A combined edge and corner detector." In Proc. Alvey Vision Conference, pp.147–151, 1988.
9. R. Klette. "Concise Computer Vision." Springer, London, 2014
10. C. B. Suo, D. Q. Yang, Y. P. Liu. "Comparing SIFT, SURF, BRISK, ORB and FREAK in Some Different Perspectives[J]."Beijing Surveying & Mapping, 2014.
11. D. G. Lowe, "Object recognition from local scale-invariant features." In Proc. ICCV, vol.2, pp.1150–1157, 1999.
12. P M Panchal, S R Panchal, S K Shah, "A Comparison of SIFT and SURF". IJIRCCE,vol. 1, Issue 2, April 2013.
13. Pranav Gupta, Muskan Gupta, Priyanka Gulhane, "Evaluation of Different Feature Descriptor Algorithms on Classification Task", 2019 5th International Conference on Computing Communication Control and Automation (ICCUBEA), 978-1-7281-4042-1/19/$31.00 ©2019 IEEE, 2019.
14. Cai-Wei Shu, and Xue-Zhong Xiao, "ORB-Oriented Mismatching Feature Points Elimination". 978-1-5386-7672-1/18/$31.00 ©2018 IEEE, 2018.
15. Jinjin Yao12, Pengchao Zhang12, Yan Wang12, Zhaoyang Luo 12, Xiaohui, "An Adaptive Uniform Distribution ORB Based on Improved Quadtree", Digital Object Identifier, 10.1109/ACCESS, 2017.
16. Shuo Li and Zhiqiang Wang, "A Research of ORB Feature Matching Algorithm Based on Fusion Descriptor", 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020), 978-1-7281-4323-1/20/$31.00 ©2020 IEEE, 2010.
17. Maosen Wang ,Shaozhang Niu,Xuan Yang, "A novel panoramic image stitching algorithm based on ORB", Proceedings of the 2017 IEEE International Conference on Applied System Innovation IEEE-ICASI 2017 - Meen, Prior & Lam (Eds) 818 - ISBN 978-1-5090-4897-7, 2017.