iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** II **Month of publication:** February 2022

DOI: https://doi.org/10.22214/ijraset.2022.40395

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

Fast and Reliable Vision-Based Navigation for Real Time Kinematic Applications

Mustafa M. Amami
Civil Engineering Department, Benghazi University

Abstract: ***Automatic Image Matching (AIM) is the term used to identify the automatic detection of corresponding points located on the overlapping areas of multiple images. AIM is extensively used with Mobile Mapping System (MMS) for different engineering applications, such as highway infrastructure mapping, monitoring of road surface quality and markings, telecommunication, emergency response, and collecting data for Geographical Information Systems (GIS). Robotics community and Simultaneous Localization And Mapping (SLAM) based applications are other important areas that require fact and well-distributed AIM for robust vision navigation solutions. Different robust feature detection methods are commonly used for AIM, such as Scale Invariant Feature Transform (SIFT), Principal Component Analysis (PCA)–SIFT and Speeded Up Robust Features (SURF). The performance of such techniques have been widely investigated and compared, showing high capability to provide reliable and precise results. However, these techniques are still limited to be used for real and nearly real time SLAM based applications, such as intelligent Robots and low-cost Unmanned Aircraft Vehicles (UAV) based on vision navigation. The main limitations of these AIM techniques are represented in the relatively long processing time and the random distribution of matched points over the common area between images. This paper works on overcoming these two limitations, providing extremely fast AIM with well- distributed common points for robust real time vision navigation. Digital image pyramid, Epipolar line and 2D transformation have been utilized for limiting the size of search windows significantly and determining the rotating angle and scale level of features, reducing the overall processing time considerably. Using limited number of well-distributed common points has also helped to speed up the automatic matching besides providing robust vision navigation solution. The idea has been tested with terrestrial MMS images, and surveying UAV aerial images. The results reflect the high capability of the followed technique in providing fast and robust AIM for real-time SLAM based applications.***

Keywords: Automatic Image Matching, Epipolar Line, Image Pyramid, SLAM, Vision Navigation, Real Time, Vision Navigation.

I. INTRODUCTION & CRITICAL REVIEW

AIM can be defined as the automatic detection of corresponding points on the overlapping areas of multiple images. AIM is extensively used with MMS for a great deal of engineering applications, such as highway infrastructure mapping, monitoring of road surface quality and markings, telecommunication, emergency response, and collect data for GIS [1]. Robotics community and SLAM based applications are other important areas that require precise and well-distributed AIM for robust vision navigation solutions [2]-[4]. Several different methods have been introduced for AIM but, in general, these methods can be classified into three categories, namely: area-based matching, feature-based matching, and relation-based matching. The most common techniques used with these methods are cross-correlation and least squares correlation. Finding out the corresponding points between two overlapped images is a part of many photogrammetry and computer vision applications, such as image registration, camera calibration, object detection, and image recovery. AIM, in general, tends to include three main steps. In the first, the interest points in each image are selected at distinctive locations, such as corners, edges, blobs, and T-junctions. Repeatability is regarded as the most important property of any an interest point detector, where it reflects the detector reliability for finding the same physical interest points under different viewing conditions. After that, the neighbourhood of each selected interest point is represented by a feature vector. For successful matching, the descriptor should be distinctive and robust to: noise, detection displacements and geometric and photometric deformations. In the last step, the calculated descriptor vectors of the different images are matched based on a distance between the vectors, e.g. Euclidean distance. The dimension of the descriptor plays a significant role in the processing time where less dimensions are recommended for faster image matching. However, using lower dimensional feature vectors usually leads to less distinctive descriptor, resulting degraded outputs. For more details, the reader is referred to [4]-[8].

During the last a few years, the integration of Inertial Navigation System (INS) and vision navigation has started to be used for reducing the cost of MMS and overcoming the main dependency of MMS on Global Positioning System (GPS) [9], [10].

©IJRASET: All Rights are Reserved

Such integration is used widely in robotics community for applying the concept of SLAM, which can be used for image based navigation and integrated with low-cost GPS solutions [11] and might be tested with different types of GPS antennas to overcome some limitations [12]. SLAM can be defined as the procedure in which a map is created and used simultaneously for location. In other words, the robot position should be known from the map to draw the map. Therefore, positioning is simultaneously solved by sequential localization and mapping. The idea of vision navigation used in some MMS is similar to SLAM in terms of finding out the camera Exterior Orientation Parameters (EOP) from the images (resection) and using these parameters for creating more points on the images (intersection) which are used for calculating the camera EOPs of the following station. The simplicity of vision navigation technique used in MMS comes from the fact that in MMS, the real time processing may not be unnecessary as in the case of a robot. This helps to bundle the observations of several images together providing better and more robust results. The concept of vision navigation and SLAM is illustrated in figure (1) for the case of a pair of cameras. In the first station, the system is localized based on the relative displacement with respect to some common points (red targets) between the two cameras (resection). After that, the relative position of more points (circular targets) located in the overlap area between the two images can be determined (intersection). In the second station, where the system moves to another position, the already determined targets in the previous station are used to locate the system (resection). Then, the positions of more points are determined (intersection) which are used in the following station, and so on. Theoretically, the vision navigation can be carried out in relative coordinate system and when necessary, three Ground Control Points (GCPs) can be used for absolute positioning. However, this is often hard to be attained practically and an additional sensor, such as Inertial Measurement Unite (IMU) is often employed to provide initial values for the camera positions and rotations for more reliable and dependable bundle adjustment solution. The additional sensor can also be used to fill in the gaps when the vision navigation is failed. For more details, the reader is referred to: [2], [11], [12], [13], and [14].

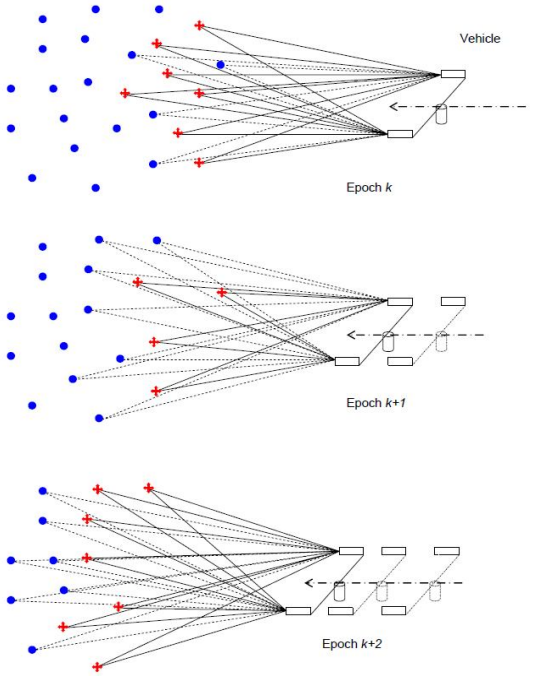


Fig. 1 SLAM concept [13]

Scale Invariant Feature Transform (SIFT), Principal Component Analysis (PCA)–SIFT and Speeded Up Robust Features (SURF) are considered to be from the most common automatic image matching methods used in photogrammetry and computer vision applications, such as SLAM, vision based navigation, image registration, camera calibration, automatic image mosaic, indexing, recognizing panoramas, and traffic sign recognition. These AIM algorithms consist generally of two processes: feature point detection and description. The first aims to find the interest points and should be robust to rotation, scaling and image noise, whereas the second is to construct unique distinctive descriptors for the feature points on the first image to be reliably identified from those on the other image [15].

SIFT were introduced by David Lowe in 2004 as a scale space based feature matching technique. The algorithm is regarded as a powerful tool in the area of automatic images matching, with high ability to extract stable features. SIFT is one of the most robust local invariant detector and descriptor with respect to geometrical changes. SIFT was designed to be invariance to image scale, rotation and affine transformation which might attribute the wide spared of this technique in photogrammetry and computer vision applications, such as image mapping, recognition, 3D modelling, GIS, and vision navigation. Four main steps are considered in SIFT, namely: the extreme point detection in scale space, precise positioning of interest points, assigning the main orientation of these points, and constructing a unique distinctive descriptor for each interest point. For faster processing time, Difference-of-Gaussian (DOG) is used in SIFT instead of Gaussian in the first step to find out possible interest scale and orientation invariant points [16], [6].

For faster matching and considerable space advantages, PCA has been used in PCA-SIFT to normalize gradient patch as an alternative for histograms. PCA is a dimensionality reduction method exploited in PCA-SIFT to make the feature vector considerably smaller than that of SIFT. This helps to reduce computations and as a consequence decrease the processing time and save significant storage space [17]. SURF has been developed to overcome the limitations of image matching algorithms, such as SIFT and PCA-SIFT, in terms of processing time using an intermediate image representation known as integral image and Fast-Hessian detector. This intermediate image can be computed rapidly from an input image as the summation of the intensity values between the point and the origin. This can help to speed up the time of any upright rectangular calculation considerably. Detection procedures in SIFT and SURF are different to some extends. SIFT creates an image layers which are filtered individually with Gaussians and uses the difference of sigma values. This is not the case with SURF where a "stack" is generated providing images of the same size. Using the integral images in SURF helps to filter the stack using a box filter approximation of second-order Gaussian partial derivatives as the computation of rectangular box filters can be applied in near constant time [5], [18].

Different researches, such as [3], [19], and [20] have investigated and compared the performance of such matching methods for image deformation, such as scale changes, rotation, blur, compression, illumination changes, and affine transformations. The performance of the three robust feature detection methods adopted, namely SIFT, PCA-SIFT, and SURF has been investigated and compared. K-Nearest Neighbour and Random Sample Consensus (RANSAC) have been adopted for evaluating and analysing the results. K-Nearest Neighbour has been utilized for getting the common points, which are filtered using RANSAC to determine the number of correct matched points. The performance has been assessed using the repeatability where higher repeatability means more stability. For evaluating the accuracy, the number of correct matched points achieved from RANSAC has been used. For reliable investigation, the same image dataset as well as PC and operating system have been used. Processing time including feature detection, description and matching has been determined for evaluating the fastness of each method. Results show that SURF is relatively the fastest comparing to SIFT and PCA-SIFT, which is attributed to utilizing integral image. Also, SURF detector known as 'Fast-Hessian' is more than three times faster than SIFT detector and five times faster than Hessian-Laplace. SIFT shows high stability with most image deformation cases although it's slow. PCA-SIFT is faster than SIFT to some extends with good stability in rotation and illumination changes.

Testing SURF shows that the performances of Hessian matrix based detector and adopted descriptor are comparable to the interest point detectors and descriptors used with conventional matching algorithms, such as SIFT, especially with normal rotations and scales. The high repeatability of the SURF detector is advantageous for different applications, such as camera self-calibration where the accuracy of the interest point has a direct effect on the whole Self Calibration Bundle Adjustment Solution (SCBAS). The detector speed is another important advantage of SURF over SIFT which is useful for real-time computation. Also, using integral images make SURF descriptor competitive in terms of processing time which can allow using longer victors (more surrounded points) for describing the interest points. SURF shows high performance with camera calibration and object recognition where the accuracy of the interest points and the descriptor distinctiveness play a considerable role on obtaining more accurate 3D reconstruction [5].

In [3], the author has worked on enhancing the performance of common AIM methods, namely SIRF, PCA-SIRF and SURF in terms of processing time using image pyramid. The image pyramid is a data structure showing the same image with different resolution rates. Figure 1 shows image pyramid with different resolutions. There are different resampling methods available for generating an image pyramid, such as nearest neighbour, bilinear interpolation and bi-cubic interpolation. In nearest neighbour method, pixels on the resampled image take the intensity value of the pixels on the original image within which the points fall. In bilinear interpolation and bi-cubic interpolation methods, the intensity value of the resampled pixel equals to the weighted average of the intensity value of the original pixels in the nearest 2 by 2 and 4 by 4 area, respectively. The images are firstly resampled and matched to detect the interest points.

Then, the approximate locations of the matched points are determined on the original images from similar triangles. These points are surrounded by small searching windows and matched again with the corresponding searching windows in the other image. As a result, instead of matching the whole two images, a number of tiny images are matched together. Different tests were carried out for evaluating the processing time of each algorithm with and without applying the image pyramid with different image resolutions, different image resampling levels and techniques, and different numbers of matched points.

The results showed that the applied idea is powerful in terms of reducing the processing time. The performance of utilizing image pyramid is affected by the resampling level where the higher the rate, the faster the processing time. However, higher resampling levels have an effect on the number of detected points, especially with non-interpolation methods, such as nearest neighbour. The resampling method has also an effect on the processing time of the introduced idea and this effect increases with increasing the image resolutions. The idea worked well with high image resolutions where the differences between the obtained processing time and the original time are significant. The results showed also that the processing time is affected by the number of required matching points where the smaller the number, the faster the technique. In general, the suggested idea for speeding up SIRF, PCA-SIRF and SURF was effective in reducing the processing time considerably. Although each method of AIM technique has its own advantages, and although the serious efforts have been applied in terms of speeding up the required processing time towards real time applications, these techniques are still considered to be slow, especially for fast real time applications, such as vision navigation based robots and low-cost surveying UAVs [3].

As for the Epipolar line or Epipolar plane used in this paper, it is that plane pass throughout four points, namely: the exposure stations of the two cameras and the images of any feature on the two photos. In co-planarity equation, the EOPs of the first image can be fixed as zeros, which means the same image coordinates and camera focal length can be used directly in the equation. Those of the second image should be converted to be parallel to the first image using the relative rotations between the two images. The co-planarity equation can be written as:

$$BY\,(xa1\ za2 - za1\ xa2) + BX\,(ya2\ za1 - ya1\ za2) + BZ\,(ya1\ xa2 - ya2\ xa1) = 0 \qquad (1)$$

Where,

BX, BY, BZ are differences between the camera capturing centres in three directions.

xa1 , ya1, za1 are the transformed image (1) coordinates to be parallel to the absolute coordinate system.

xa2 , ya2, za2 are the transformed image (2) coordinates to be parallel to the absolute coordinate system.

xa1 , ya1, za1 are functions of three rotations of the first image (ω1, ρ1 and κ1).

xa2 , ya2, za2 are functions of three rotations of the second image (ω2, ρ2 and κ2).

Epipolar line is effective in Y direction, but not in X direction where mismatched points cannot be detected. This is clear from the co-planarity equation where the effect of X component decreases when the difference between Y components in the two images rounds about zero. This is also the case with slop images and images with significant differences in Y and X components, where mismatched points located on the lines parallel to the baseline of the two cameras may not be undetected. Figure (2) shows an example of mismatched points cannot be detected using Epipolar line filter.

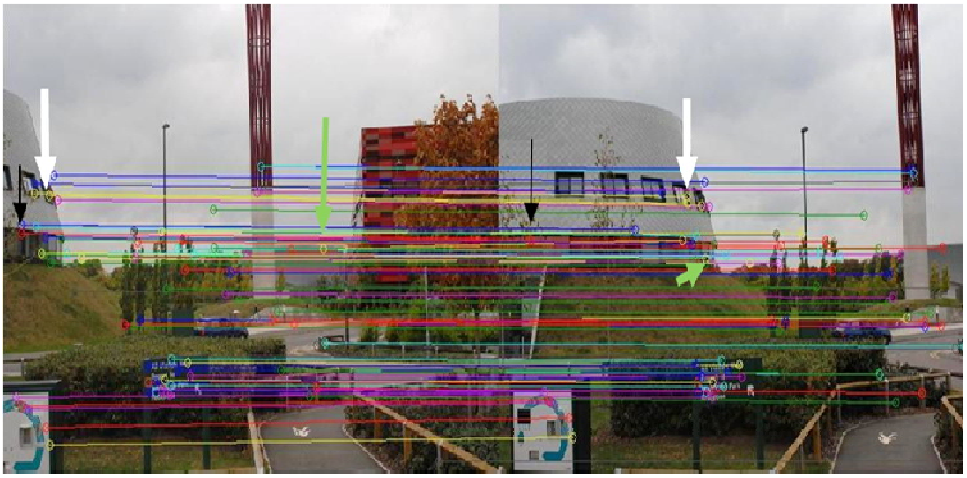


Fig. 2 Mismatched points undetected by Epipolar line filter

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

Epipolar line will be integrated with 2D transformation for detecting the mismatched points, providing robust well-distributed common points across the overlapped area. Epipolar line will be firstly used to deal with the first group of matched points until reliable parameters for 2D transformation are determined. At least two points can be used for determining the four parameters of linear 2D transformation, but more points tend to be used and solved by least squares method for reliable solution and detecting any outliers in observations. Data Snooping Method will be used in this paper due to its simplicity and affectivity, where it can deal with data including a number of outliers. However, for more reliable results, the percentage of gross errors to the whole observations should be as small as possible. The residual of each observation should not be bigger than maximum 3 times its standard deviation to pass the filter. This value can be fluctuated from 0 to 3 depending on the required confidence level. With 99% confidence level, the value is nearly 2.6 and with 99.99%, it is almost 3. When the residual overcomes the chosen threshold, the observation is considered to be an outlier with (1- confidence level) probability of rejection the observation when it should be accepted. Then, the determined four parameters of 2D transformation will be used for providing initial position on the photo for any common point detected on the other image. This can help the AIM techniques for reducing the size of search area considerably and consequently, faster and robust corresponding common points can be achieved [3, 2, 4].

For determining the 3D displacements of the camera position, the relationship between the two images and the objects common between these images should be known. This relationship can be illustrated by the co-linearity equations. Co-linearity condition specifies that, for any image, the ground point and its corresponding image point must lie along a straight line with the exposure station.

$$xa = -f\,([r11(X0 - XA) + r12\,(Y0 - YA) + r13\,(Z0 - ZA)]) / ([r31\,(X0 - XA) + r32\,(Y0 - YA) + r33\,(Z0 - ZA)]) \quad (2)$$

$$ya = -f\,([r21(X0 - XA) + r22\,(Y0 - YA) + r23\,(Z0 - ZA)]) / ([r31\,(X0 - XA) + r32\,(Y0 - YA) + r33\,(Z0 - ZA)]) \quad (3)$$

Where,

xa and ya	*are the image coordinates of point A.*
X0, Y0, Z0 and XA, YA, ZA	*are the coordinates of capturing centre and object A.*
r	*is a rotation matrix including three rotations of each image about X, Y and Z axes.*
f	*is the camera focal length.*

This paper works on speeding up AIM techniques using Epipolar line and 2D transformation beside image pyramid for providing fast and robust navigation solution for real time SLAM based applications. The works in this paper is considered as an extension for the previous efforts carried out by [3] for speeding up AIM using just image pyramid. The idea will be tested with three famous AIM methods, namely: SIRF, PCA-SIRF and SURF. The idea depends mainly on limiting the size of search windows significantly and determining the relative rotation and scale factor of features, which helps on reducing the overall processing time considerably. Using limited number of well-distributed common points will also be tested to speed up the automatic matching beside provide robust vision navigation solution. The idea will be tested with terrestrial images taken from low-cost vision based personal MMS designed by the author. See figure (3). This is in addition to aerial images captured by low-cost surveying UAV in 2020 at Benghazi University.



Fig. 3 Right: low-cost vision base personal MMS & left: UAV site works

©IJRASET: All Rights are Reserved

II. METHODOLOGY

The methodology followed in this paper can be illustrated in the following steps:

1) Determining the parameters of the Epipolar line between the two images. This can be done by finding out the centre of each image on the other image using normal AIM technique. The line pass throughout the four points is the Epipolar line. See figure (4). The four points make 8 equations in just 4 unknowns, which means least squares should be used for the solution. The residuals of the 8 observations can indicate the quality of the matched two points, where the residuals should be in the level of one pixel.

2) Determining the value of relative feature rotation and scale factor. The X axis of the two images may not be parallel or located on the same line, which means the right image is relatively oriented comparing to the left image. The value of this orientation can be determined from the coordinates of each image centre on the other image. See figure (4). This value is extremely useful where the range of rotations will be reduced significantly. This means that instead of carrying out the matching procedure with different rotations (tend to start from -45o to 45o), the rotation angle will be nearly known and it may be useful to give a flexibility range of nearly +/- 5o to increase the confidence level. The scale is another factor can be estimated in this step, where the percentage between the lengths of the same distances on the two images gives a clear indication about the scale factor. This is also important for AIM techniques which are based on matching the interest points with different scaling levels. However, the estimated value of scale factor should also be flexible where the scale in photos is different from point to point across the image due to the central projection, especially with tilted photos.

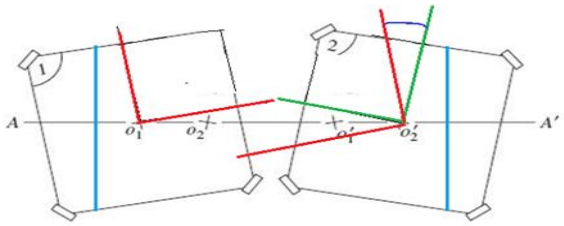


Fig. 4 Epipolar line (O1, O2', O1', O2), estimated rotation angle, & overlap area

3) Determining the area of overlapping between the two images automatically. This can help to limit the search area in AIM, reducing the processing time considerably. The last point located on the Epipolar line in the left image will be automatically matched with the Epipolar line narrow strip area on the right image to find out the corresponding point. And the same for the right image centre. The orientation and scale factor values determined in step (2) will be used in this step to reduce the search time. The border lines in Y direction will be perpendicular on the Epipolar line. See figure (4).

4) The determined overlap area between the two images is divided into a number of subareas. This number is a function of the image shape and quality. If the image is square and the captured area is rich with features, it can be divided into 16 subareas. The number of subareas is recommended to be increased in feature-less areas, where the high redundancy can help to overcome the geometrically weak points. For rectangular images, more subareas in X axis can be used and 20 points is more than enough for carrying out the concept of vision navigation and SLAM. The minimum number of common points for relative orientation is 5 if the flaying height or the distance between the two cameras is known in the Epipolar line direction (air base). If not, 6 well-distributed reliable common points are necessary to solve the 6 EOPs of the second image and the 3D relative coordinates of each point. Using 16 points, for example, gives 74 co-linearity equations in 54 unknowns (6 EOPs + 3 * 16 relative coordinates), which provides a good degree of freedom. The three Interior Orientation Elements (IOEs) of the camera can also be solved in this case as unknowns and this procedure is known in photogrammetry as self-calibration bundle block adjustment.

5) The idea of speeding up AIM, applied in [3] will then be used for finding out the most interest points in each subarea of the two images. This idea is based on the truth that there is a high probability for each interest point in the original image to stay as an interest point in the different image pyramid levels. Based on that, automatic matching between any two images can be firstly performed at the resampled images to find the interest common points between these images. This can be carried out quickly with the low resolution levels where the lower the image size, the faster the processing time. Then, the approximate locations of just two interest points on two subareas are determined on the original images based on basic similar triangles. The approximate locations of these two points on the original image are then surrounded by small searching windows and matched again with the corresponding searching windows in the second image. Epipolar line and the determined orientation and scale values will be used with image pyramid for reducing the search area and limiting the rotating and scale ranges in AIM techniques.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

6) In addition to these two points matched in step (5), there are 4 other points matched before: two in step (1) and two in step (3). These 6 points will be used to determine the four parameters of 2D transformation between the two images. 6 points makes 12 equations in 4 unknowns, giving a good chance for evaluating the residuals of observation and detecting any outliers.

7) Then, the rest of interest points in the other subareas will be matched as explained in step (5), but instead of using Epipolar line to limit the search area, the determined 2D transformation parameters will be used, giving more precise estimated positioning. This will help to reduce the search time considerably and provide error-free matched points.

8) Determining the 6 EOPs of the second image as relative values to the first image using co-linearity equations. The EOPs include 3D displacements (ΔX, ΔY, ΔZ) beside the changes in rotations (Δω, Δρ, Δκ) about (X, Y, Z), respectively. For vision navigation and SLAM based applications, just the 3D displacements tend to be used to indicate the change in position. The methodology can be summarized in the following workflow diagram.

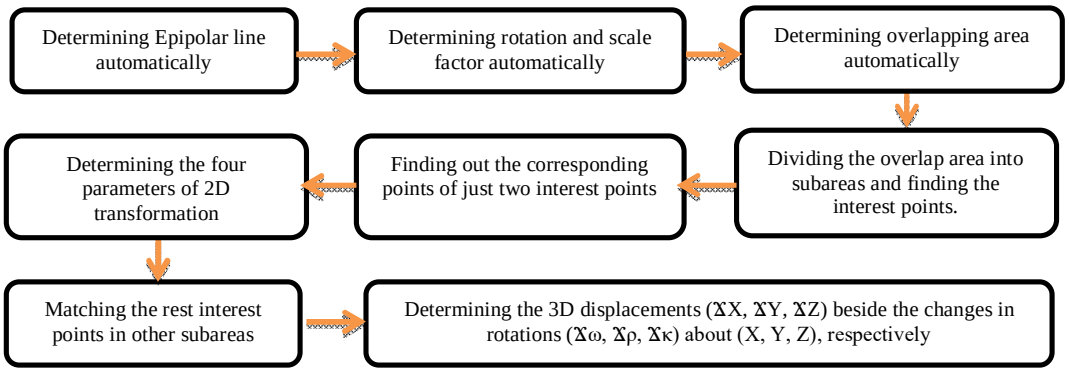


Fig. 5 Workflow diagram

III.RESULTS AND DISCUSSIONS

Three famous AIM techniques have been tested with the suggested idea in this paper, namely: SIRF, PCA-SIRF and SURF. The first test have be applied to evaluate the advantages of new idea over the previous one introduced by [3]. Two 10 megapixel resolution terrestrial images have been used in this test with the settings of 25% resampling rate, 9 required matched points, and bi-cubic interpolation method. Figure (6) shows an example of the aerial images used in this test and well-distributed common points across the overlap area.

Table 1 illustrates the processing time in milliseconds of: the original three AIM techniques, the techniques with just image pyramid (WIP), and with the new idea (WNI) introduced in this paper. It should be mentioned that the processing time is based mainly on the computer installed memory (RAM) and in this paper, the same computer with 4.00 GB RAM has been used for all tests. These days, high speed computers with RAMs in the range of 16 GB and more are used for smart robots and aerial vision based navigation.

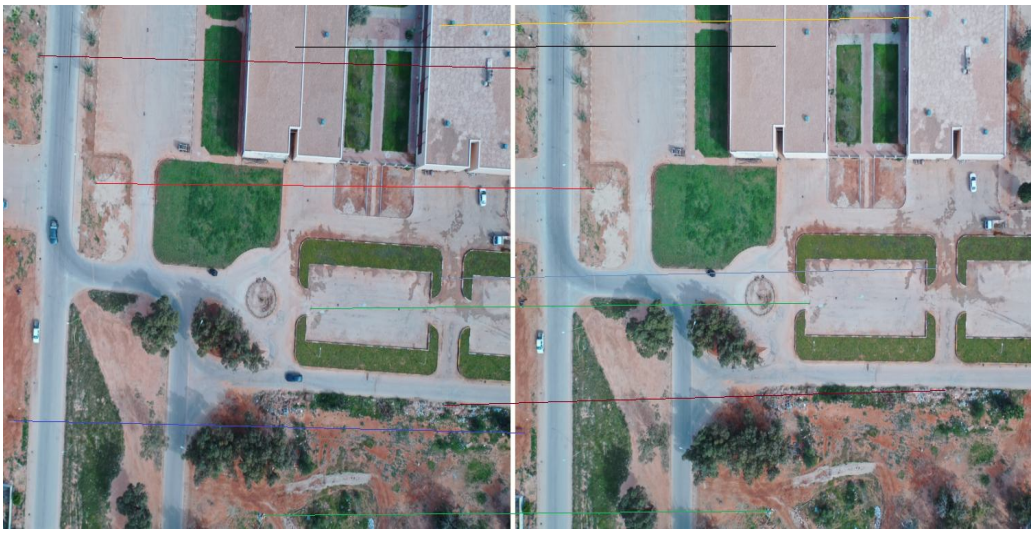


Fig.6 Well-distributed robust automatic matched points

TABLE 1
SIFT, PCA-SIFT, AND SURF PROCESSING TIME FOR 9 POINTS (MILLISECONDS)

SIFT			PCA-SIFT			SURF		
WIP	Original	WNI	WIP	Original	WNI	WIP	Original	WNI
33381	163568	9537	28331	152988	8332	6930	47819	2163
WIP	WIN /WIP	WNI	WIP	WIN /WIP	WNI	WIP	WIN /WIP	WNI
20.4%	28%	5.8%	18.5%	29%	5.4%	14.5%	31%	4.5%

From the results, it is clear that with the suggested idea, the processing time of SIRF, PCA-SIRF and SURF is improved considerably comparing to the original time and also to that enhanced with image pyramid only. The order of the three algorithms in terms of processing time is the same before and after applying the suggested technique, where SURF and SIFT are the fastest and slowest, respectively. The results show also that SURF has the highest improvement rate (lowest processing time), which can be attributed to the truth that the suggested idea depends on repeating the algorithm by a number of times equals to the selected matching points + 4 (2 for determining the center of each image in the other, and two for determining the overlap area). This means that with 9 required matching points, SURF, for example, will be used 13 times to get the final matched common points. As a result, the faster the algorithm without the suggested idea, the higher processing time improvement rate can be achieved with this suggested idea. However, when comparing the changes in the two rates, it can be seen that the new techniques helps SIFT with nearly the same level of PCA-SIFT, and SURF.

This can be attributed to the fact that SIFT is designed to deal with wide range of rotations and scales. Therefore, when providing SIFT with initial rotation and scale, searching for the corresponding feature will be limited and this reflects significantly on the whole processing time. It should be mentioned that it is difficult to use the processing time and the improvement rates shown in the table as absolute and fixed values, where they depend on the number of matching points, resample level, resample method, and original image size. The second test has been performed to investigate the effect of resampling rate on the suggested technique. Three different resampling levels (25, 50, & 75) have been used with the settings of bicubic interpolation method and 9 matching points. Table 2 illustrates the processing time of each resampling level for each AIM technique comparing to that based just on image pyramid.

TABLE 2
PROCESSING TIME OF SIFT, PCA-SIFT, AND SURF WITH DIFFERENT RESAMPLING RATES (MILLISECONDS)

SIFT (Image Pyramid)			PCA-SIFT (Image Pyramid)			SURF (Image Pyramid)		
25%	50%	75%	25%	50%	75%	25%	50%	75%
33381	55367	117513	28331	49995	102880	6930	11780	29630
SIFT (New Idea)			PCA-SIFT (New Idea)			SURF (New Idea)		
25%	50%	75%	25%	50%	75%	25%	50%	75%
9537	14980	32376	8332	13618	29834	2163	3657	8951

It is clear from the results that the higher the resampling rate, the faster the processing time. This is theoretically expected as high resample level decreases the large amount of image data and as a consequence, the computation time is reduced significantly. As the suggested technique is based on applying the matching algorithm on the resampled images firstly before reapplying the algorithm on the resulted tiny images, the resampling level will play a significant role in reducing the processing time as clear from the table. The results show the high capability of the suggested technique in speeding up the AIM processing time, especially with high image resampling rates. SURF has the fastest processing time for all resample levels, PCA-SIFT comes second and SIFT is last. The effect of image resolution on the suggested technique has also been investigated using three different resolution aerial images captured by low-cost surveying UAV. The number of matching points, resampling rate, and resampling method have been fixed as 12 points, 20%, and bi-cubic interpolation, respectively and images with resolutions of 10, 12, and 20 megapixel have been used. Theoretically, this effect is similar to changing the resampling rate, but in different way where the higher the image resolution, the slower the processing time. As seen from Table 3, the difference between the speeded up processing time and the original and image pyramid based time becomes more and more significant with increasing the image resolution. This can be attributed to utilizing Epipolar line and 2D transformation, beside image pyramid, for reducing the search areas to great extends and finding out the estimated value of rotating and scaling between images.

TABLE 3
SIFT, PCA-SIFT, & SURF PROCESSING TIME WITH DIFFERENT IMAGE RESOLUTIONS (MILLISECONDS)

SIFT (Image Pyramid)			PCA-SIFT (Image Pyramid)			SURF (Image Pyramid)		
10 MB	12 MB	20 MB	10 MB	12 MB	20 MB	10 MB	12 MB	20 MB
26817	31879	39680	23545	29245	35924	5543	6489	8581
SIFT (New Idea)			PCA-SIFT (New Idea)			SURF (New Idea)		
7556	8795	10972	6798	8117	10235	1597	1743	2301

The last test has been applied to study the effect of changing the resampling methods, namely: nearest neighbour, bilinear interpolation and bi-cubic interpolation on the performance of the suggested idea. The same resampling rate has been chosen to match the same number of points with the three AIM techniques. The results show that the differences in processing time between the three resample methods are as small as can be neglected. This can be attributed to the fact that in the suggested technique used in this paper, the resampling procedure is not repeatable and it is used once at the beginning when resampling the image. Nearest neighbor provides relatively the fastest processing time and bi-cubic interpolation is the slowest which is attributable to the differences in the amount of calculations between the methods. The difference in processing time between the three resampling methods can become more significant with higher resolution images where more calculations are required by the interpolation resampling techniques. The interpolation methods are theoretically better than the nearest neighbour method where the intensity values of all pixels are utilized allowing all interest points in the original image to be considered in the resampled image. This is not the case with the nearest neighbour method where the resulted pixel takes just the intensity value of the corresponding pixel in the original image and consequently, some interest points might be neglected. With high resample rates, the differences in the performances between nearest neighbour and the other interpolation based methods become significant in terms of detecting high number of interest points where lots of these points may not be considered using nearest neighbor. Reducing the number of detected interest points may not be suitable for some engineering applications that require dins matched point cloud, such as creating digital elevation models and ortho-images. On the other hand, there are many subjects that need fast processing time for real or nearly real time applications with limited number of matching points, such as vision navigation, robots, and SLAM.

Finally, the suggested technique has been used for determining the EOPs of 12 pairs of photos taken by low-cost vision based personal MMS designed by the author in 2014, at Nottingham university as a part of his PhD. 5 high resolution aerial images taken by low-cost surveying UAV have also been processed to determine the 3D displacements in the camera exposure stations. To evaluate the quality of results, the precise camera positions have been determined using well-distributed GCPs in both cases. Figure (7) shows 3D simulation of the terrestrial camera trajectory beside the true trajectory determined by vision navigation. Table (4) illustrates the Roth Mean Square Error of the EOPs and a number of check points.

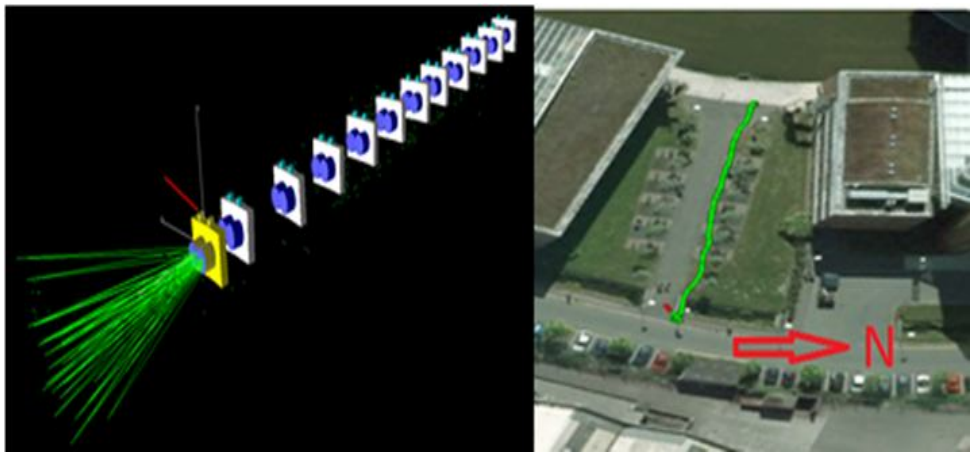


Fig. 7 3D simulation of the terrestrial camera trajectory beside the true trajectory

TABLE 4
EVALUATING THE PERFORMANCE OF THE SUGGESTED TECHNIQUE IN SLAM BASED APPLICATIONS

Check points RMSE (Terrestrial)	X m		Y m		Z m	
	0.063		0.041		0.033	
EOPs RMSE (Terrestrial)	X0 m	Y0 m	Z0 m	*w* deg.	*p* deg.	*k* deg.
	0.032	0.024	0.022	0.340	0.32	0.26
Check points RMSE (Aerial)	X m		Y m		Z m	
	0.042		0.037		0.071	
EOPs RMSE (Aerial)	X0 m	Y0 m	Z0 m	*w* deg.	*p* deg.	*k* deg.
	0.025	0.030	0.041	0.31	0.42	0.46

It is clear from the table that the camera optical axes (X axis in the case of terrestrial images and Z axis in the case of aerial photos) have relatively the lowest precision level. This is attributed to the fact that when image coordinates are measured with small errors, the two lines passing from the exposure centres of the two cameras to the point will intersect at a point on a distance (s) from the true position. See figure (8). This error is a function of the intersection angle and the distance between the camera and the object, where the farther the object, the worse the quality. This means that when the distance between cameras is small, just close objects can be used for acceptable results. However, the results show the ability of the introduced idea to be used successfully in SLAM, providing fast and robust vision navigation with 3D quality of less than 1 decimetre.

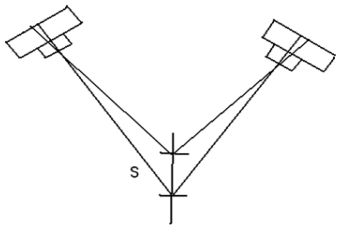


Fig. 8 Intersection angle effect

IV. CONCLUSION

This paper has worked on overcoming two of the main AIM limitations, namely the long processing time and the random distribution of matched points. The works in this paper is considered as an extension for the previous efforts carried out by the author for speeding up AIM using just image pyramid. In this paper, digital image pyramid, Epipolar line and 2D transformation have been utilized together for limiting the size of search windows significantly and determining the general rotation and scale of features, reducing the overall processing time considerably. Using limited number of well-distributed common points has also helped to speed up the automatic matching besides providing robust vision navigation solution. Different tests have been carried out for evaluating the suggested idea with SIRF, PCA-SIRF and SURF. The evaluation has included using different image resolutions, different image resampling levels and techniques. The results show that the introduced idea is really powerful where the processing time has reduced to be (4.5%, 5.4%, and 5.8%) from the original time for SIRF, PCA-SIRF and SURF, respectively. The processing time has also reduced by nearly 30% comparing to utilizing just the image pyramid. The performance of the suggested method is affected by the resampling level, where the higher the rate, the faster the processing time. However, higher resampling levels might have an effect on the number of detected points, especially with non-interpolation methods, such as nearest neighbor. The resampling method has also an effect on the processing time of the introduced idea and this effect increases with increasing the image resolutions. The idea works well with high image resolutions where the differences between the obtained processing time and the original time are significant. In general, the results have reflected the ability of the introduced idea to be used successfully in SLAM, providing fast and robust vision navigation with 3D quality of less than 1 decimeter. This might be extremely useful for real time applications based on AIM, such as vision navigation based robots and real time indoor navigation.

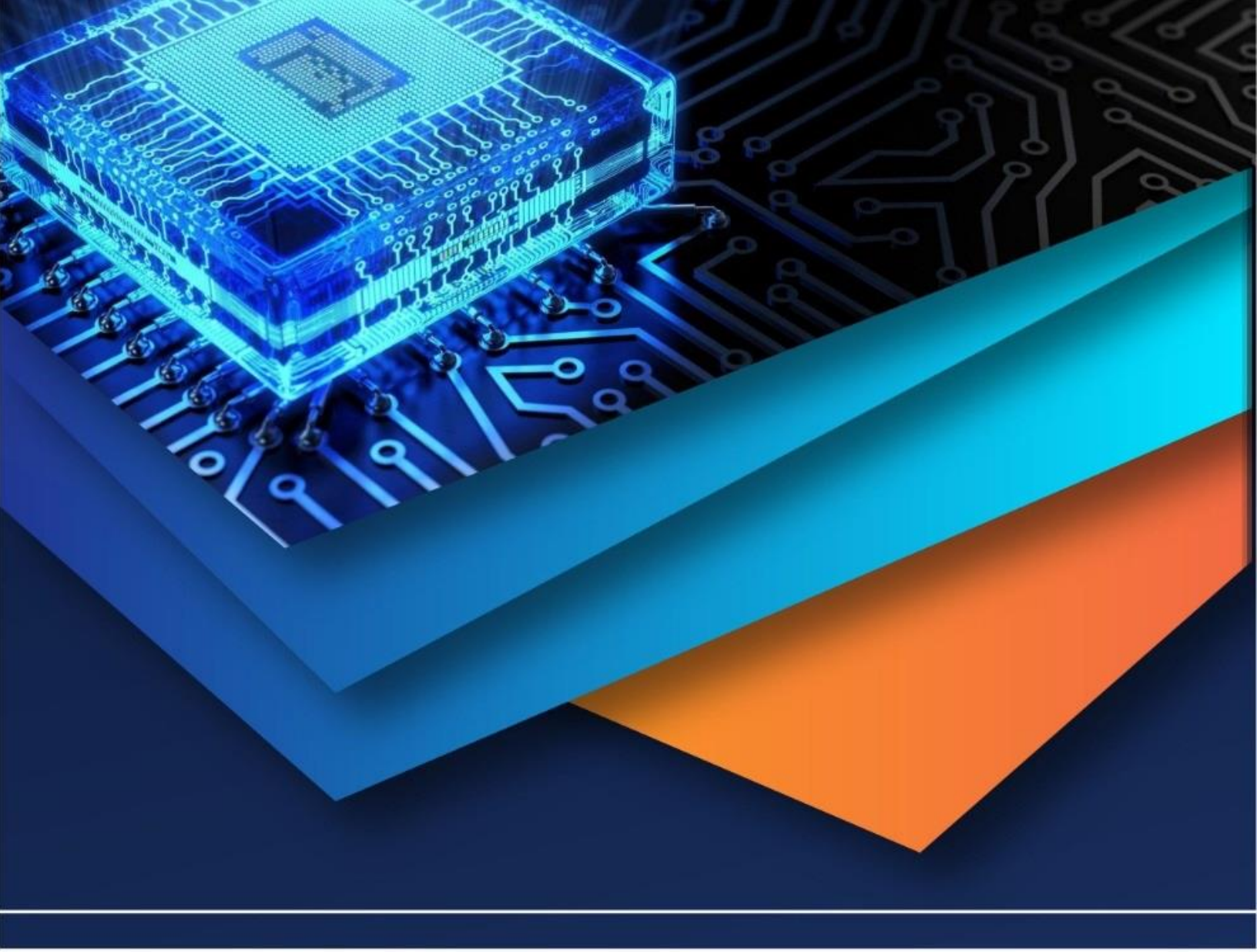
V. ACKNOWLEDGMENT

I'd like to thank everyone who contribute to the implementation of this work in Benghazi University as well as Nottingham University.

REFERENCES

[1] M. C. Kus, M. Gokmen, and S. Etaner-Uyar, "Traffic Sign Recognition Using Scale Invariant Feature Transform and Color Classification,". ISCIS, vol. 8. 2003.
[2] M. M. Amami, "Low Cost Vision Based Personal Mobile Mapping System," PhD thesis, University of Nottingham, UK. 2015.
[3] M. M. Amami, "Speeding up SIFT, PCA-SIFT & SURF Using Image Pyramid," Journal of Duhok University, [S.I], vol. 20, July. 2017.
[4] M. M. Amami, M. J. Smith, and N. Kokkas, "Low Cost Vision Based Personal Mobile Mapping System," ISPRS- International Archives of The Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-3/W1, pp. 1-6. 2014
[5] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," 9th European Conference on Computer Vision. 2006.
[6] D. Lowe, "Distinctive Image Features from Scale-Invariant Key points," IJCV, vol. 60, pp. 91–110. 2004.
[7] M. Brown, and D. Lowe, "Recognizing Panoramas," Proc. Ninth Int'l Conf. Computer Vision, pp. 1218-1227. 2008.
[8] J. C. McGlone, E. M. Mikhail, J. Bethel and M. Roy. Manual of photogrammetry. 5th ed., Bethesda, Md.: American Society of Photogrammetry and Remote Sensing. 2004.
[9] M. M. Amami, "The Integration of Time-Based Single Frequency Double Differencing Carrier Phase GPS / Micro-Elctromechanical System-Based INS," International Journal of Recent Advances and technology, vol. 5, pp. 43-56. Dec. 2018.
[10] M. M. Amami, "The Advantages and Limitations of Low-Cost Single Frequency GPS/MEMS-Based INS Integration," Global Journal of Engineering and Technology Advances, vol. 10, pp. 018-031, Feb. 2022.
[11] M. M. Amami, "Enhancing Stand-Alone GPS Code Positioning Using Stand-Alone Double Differencing Carrier Phase Relative Positioning," Journal of Duhok University (Pure and Eng. Sciences), vol. 20, pp. 347-355, July. 2017.
[12] Amami, M. Testing GPS antennas with/without Frames.
[13] F. A. Bayoud, "Development of a Robotic Mobile Mapping System by Vision-Aided Inertial Navigation: A Geomatics Approach," PhD thesis, University of Calgary, Canada. 2006.
[14] H. Durrant-Whyte, and T. Bailey, "Simultaneous Localization And Mapping: Part I," Robotics and Automation Magazine, IEEE, vol. 13, pp. 99-110. 2006.
[15] A. I. Mourikis, N. Trawny, S. I. Roumeliotis, A. E. Johnson, and L. Matthies, "Vision-aided inertial navigation for precise planetary landing: Analysis and experiments,". In Proceedings of Robotics: Science and Systems, Atlanta. 2007.
[16] L. Juan, and O. A. Gwun, "Comparison of SIFT, PCA-SIFT and SURF," International Journal of Image Processing (IJIP), vol. 3, pp. 143-152. 2009.
[17] K. Peng, X. Chen, D. Zhou, and Y. Liu, "3D reconstruction based on SIFT and Harris feature points,". Proceedings of the IEEE International Conference on Robotics and Biomimetics, pp. 960-964. 2009.
[18] Y. Ke, and R. Sukthankar, "PCA-SIFT. A More Distinctive Representation for Local Image Descriptors," Proc. Conf. Computer Vision and Pattern Recognition, pp:511-517. 2004.
[19] Zhan-long Yang and Bao-long Guo, "Image mosaic based on SIFT," International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 1422-1425. 2008.
[20] P. M. Panchal, S. R. Panchal, and S. K. Shah, "A Comparison of SIFT and SURF," International Journal of Innovative Research In Computer And Communication Engineering, vol. 1, pp. 323-327. 2013.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)