International Journal on Recent Technologies in Mechanical and Electrical Engineering (IJRMEE) ISSN: 2349-7947
Volume: 2 Issue: 9 024 - 031

Mobile Robot Vision for Object Tracking using Modified SIFT

Shivanand J. Upari
Dept of Electronics & Communication Engg.
GIT College, Belagavi
Karnataka, India
E-Mail: shivanand424@gmail.com

Abstract- For a Mobile robot vision object tracking is the important application task. The main problem which has been faced in Image Based Visual Servo (IBVS) is that to detect and track objects in a series of images. In this present paper, we are going to implement a Modified SIFT algorithm to detect & track the object in various conditions such as changes in intensity, scaling factor and circular rotation. This algorithm first of all find all local feature keypoints from the image taken by the robot while is moving. In the next step, it compares these local features with the reference template. At the end based upon matched local keypoints, it approximates the position of the object, so that it can easily track the object.

*****

I. Introduction

In computer-based machines the application of vision is one that performs image processing and translation, is also known as machine vision or computer vision. In computer vision field tracking of visual object plays a great impact. It is also a division of engineering that consists of engineering in the field of mechanical and computer science, and automation in industries.

An algorithm is implemented to capture moving target positions from the stream of videos. Thus to obtain interest points of an object and to track it, the video stream is divided into multiple scenes. From these scenes the meaningful information is obtained by adopting some of the algorithms.

The visual tracking algorithm can be done in 3 methods. They are listed as follows:

1. Feature based method.
2. Differential method.
3. Correlation method.

In Feature based method, it obtains the some of the important parameters of an image and these features are tracked by applying anyone of matching algorithm. differential method, it assigns the values based up on the computation of optical flow, for the images. In correlation method, it performs the correlation between displacements of inter image sequences.

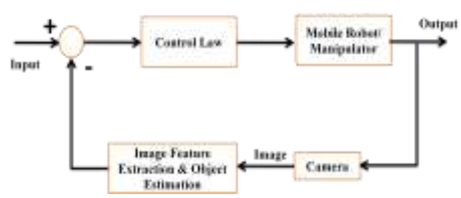

Fig. 1.1 Block diagram of Visual servoing system

The block diagram of a computer visual servoing system [11] as shown in above fig.1.1. Here the robot is going to be controlled by the control law. In this system the reference template is already stored in the database. Whenever a mobile robot captures the image by the camera, the image features extracted and these features are compared with reference template. Depends upon the matching algorithm the robot motion is controlled.

II. Literature Survey

Vision based identification of object have a great role in the commercial and industrial applications. Surveys have been focused on detecting the objects expertly by using local features keypoints for eg. pattern, size, texture, and color. During the processing of real images under different situations and surroundings, with considering methods there are number complications occurred. Some of the survey papers are explained for this project.

Haoxiang Lang, Ying Wang and Clarence W. de Silva [11] have used Scale Invariant Feature Transform (SIFT) is adopted to gain the powerful object identification &capturing the features such as enlarging, intensity changes, and revolution. This method depends mostly on symmetry of an object. Symmetry can be affected by many circumstances such as viewing angle. Major issue is in identifying the SIFT keys if the illumination levels are varying. Generally doesn't work well with lighting changes and blur.

Seelen et al. [6] have used to find front, side & rear appearance of a collection of image types by considering symmetric vertical structure with symmetry analysis & model matching. In his paper, they have stated that it must be opposed to the small variations in left and right parts of an object and also in changes in intensity. For a real operating situations this paper implements symmetry based method. But this method failed to detect the color based images and it becomes ineffective.

Buluswar and Draper [3] have adopted a method to recognition color for real-time in outdoor scenes. This technique gives the better performance in the fields such as navigation for off-road, autonomous highway and unmanned defense vehicles for target identification.

A corner-based method was proposed by ***Bertozzi*** et al. [4] to hypothesize vehicle locations. This method consists of two pipeline engines: First one, PAPRICA, a parallel architecture is used to implement for better execution of processing tasks of a lower level, and the synthesis of a particular task such as direct data I/O was enhanced. Second one, a Traditional serial architecture, used for identification of the vehicle location by executing medium level tasks. So this method was demonstrated on land vehicle in MOB LAB. This provides a strong result for hypothesizing objects by considering horizontal and vertical edges.

Matthews et al.[2] have adopted an detection of edge to search strong edges with respect vertical structure in an image. They implemented a triangular filter to restrict the right and left vehicle positions of a with respect to the vertical edge profile of the image. Even they found that the right and left positions of a vehicle by computing the highest peaks of the vertical profile.

A template based method was proposed by ***Handmann*** et al. [7]. This method implements a correlation among the template and the image. This also provides a “U” shape for identification of vehicle rear/frontal view.

A very loose template method was proposed by ***Ito*** et al. [1] to identify definite horizontal or vertical edges in actual symmetry. But this method fails to give the efficient results because of the simplicity of the template.

D. Lowe et al. [5, 10] adopted a new method for object identification and tracking which is known as Scale Invariant Feature Transform (SIFT), that uses a group of local image feature keypoints. This method gives a high reliability and robustness with respect variations scaling, noise, rotation and illumination. But this method also gives inaccurate results because it depends upon the symmetry property of the object.

In the proposed work, the Modified SIFT algorithm is implemented to achieve a mobile robot track for an object in the captured image by the camera, and gives result back to the system with respect to the information of surroundings, so that it to command the robot to find the final destination location.

III. Proposed System Implementation

In the proposed system, it is implementing an improved version of SIFT i.e Modified SIFT. This algorithm approaches the matching the feature interest keypoints of an image taken by robot camera with the template image which is stored in database. In the original SIFT algorithm it uses both local minima and maxima interest feature keypoints for image feature matching, then it evolves additional overhead for computations. Whereas in Modified SIFT it considers only maxima keypoints, that in turn reduces computational time.

The proposed Modified-SIFT employ the following major stages of computation to generate the feature descriptor of an image. For this algorithm the block diagram is as shown in fig.3.1.

The Modified SIFT algorithm has the following main stages as explained below:

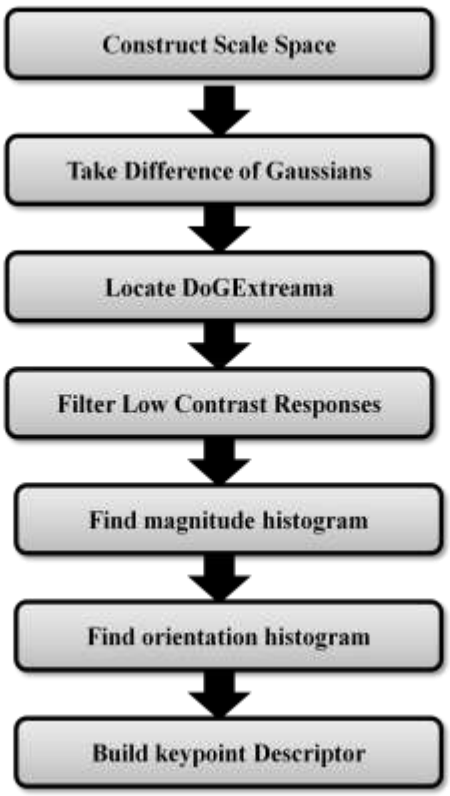

Fig 3.1: Modified-SIFT feature detection algorithm

1. **Construct Space Scale:**

First of all, the image is captured by robot camera is grouped into octaves i.e. a group of 8-rows. Next by adopting the Gaussian blur function inside an octave, successively images are blurred. Technically the convolution of an image with Gaussian operator is known as blurring.

2. **Perform Difference of Gaussian Function:**

After performing convolution, with the one scale again it has to be repeated for the other scale. By considering the two sequential images octave are chosen and subtracted from each other. Therefore this procedure once more time adopted for the next successive pair. To achieve this Gaussian function is used.

3. **Find the Extrema DOG's:**

In the previous step we have calculated the local feature keypoints. From these points now will obtain maximum keypoints, since only maximum points have the highest peak thresholds. It can be achieved by comparing the octaves with below and above.

4. **Filtration:**

This stage is mainly performs the removal of lowest insignificant feature keypoints, which doesn't carry any meaningful information. If we consider these keypoints simply it adds extra overhead for computations.

5. **Calculate magnitude and orientation histogram:**

In this stage it obtains magnitude and orientation histograms for the maximum keypoints of an image. To compute magnitude and histogram of a feature keypoints the equations (3.1) and (3.2) are used. They are listed as follows:

To find magnitude histogram,

$$\mathbf{m(x,y)} = \sqrt{\left(\boldsymbol{L(x+1,y)} - \boldsymbol{L(x-1,y)}\right)2 + \left(\boldsymbol{L(x,y+1)} - \boldsymbol{L(x,y-1)}\right)2}$$

----- (3.1)

To find orientation histogram,

$$\boldsymbol{\theta(x,y)} = \mathbf{tan}^{-1}\left(\left(\boldsymbol{L(x,y+1)} - \boldsymbol{L(x,y-1)}\right)/ \boldsymbol{L(x+1,y)} - \boldsymbol{L(x-1,y)}\right) \text{----- (3.2)}$$

6. **Construct Descriptor keypoint:**

By implementing the above formula, the descriptor keypoints are obtained for the original image captured by the robot. Here it generates descriptor for only the local maximum keypoints features, since we have already discarded the unwanted lowest peak threshold feature points.

Comparison between SIFT and Modified-SIFT algorithms:

- In Locating the kepoints, the original SIFT considers both maxima and minima, whereas modified-SIFT considers only maxima as a keypoints. This results in a lower number of keypoints being generated, which lowers the computational overhead.
- In the SIFT approach an additional overhead of removing keypoints along edges is obtained, as it requires the use of an edge detection algorithm. In the modified-SIFT approach keypoints along edges are not removed, saving this computational overhead.
- Next Modified-SIFT does not consider rotation invariance aspect of SIFT.
- Finally Creation of Modified-SIFT feature descriptor is totally different as compare to SIFT.

i.e. In an SIFT approach for calculating a feature descriptor, a 16 x 16 pixel window is used for each of the calculated maxima. Each feature is a composite histogram consisting of magnitude & orientation information.

In the modified-SIFT approach successive groups, of 8 rows of pixels, are considered, across the height of the image, to compute a feature for each group. This consists of computing histograms for magnitude and orientation separately.

Also for SIFT, total number of feature generated is very large. For each maxima, 16, 8 bin histogram i.e. 16 x 8 = 128 bins values are created, On an average in a 200 x 250 image, approximately 400 maxima will be extracted in SIFT. This willlead to 400 x 128 bins being created for comparison. Whereas in Modified-SIFT, for a 200 x 250 image (250 / 8 x 2) x 8 = 500, total bins will be created, independent of thenumber of maxima points.

IV. Results and Discussion

For the proposed project we have wirtten relevant code and that has been executed successfullly executed in MAT Lab software tool. The ouputs for each of the corresponding stages of the proposed project are explaind in breif and as shown below.

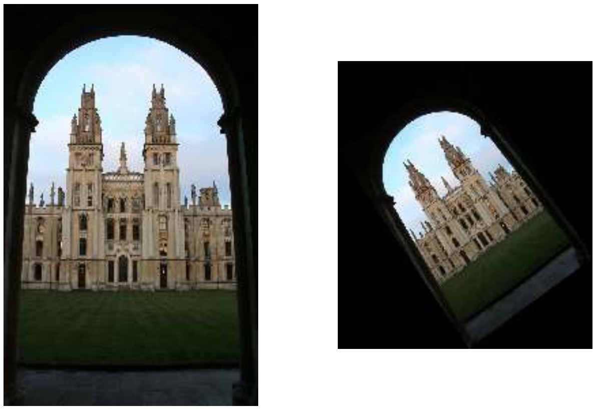

Fig 4.1: Original Image & Rotated and Scaled version

From the database the input image is taken which is already stored. The original input as it is captured by roobt, is raw one and it may or may not fit in desired scale. Therefore first of all after reading this image it has to be resized and rotated so as to remove any boundary noise. This can be shown in Fig. 4.1, it represnts input as well as resized, rotated version of it.

In the next stage will obtain the SIFT features of the refernce image which has to matched with input image. For this we will apply Modified SIFT alogrithm, so as to extract only interest keypoints as shown in Fig. 4.2, which depicts the SIFT features of reference image and resized image.

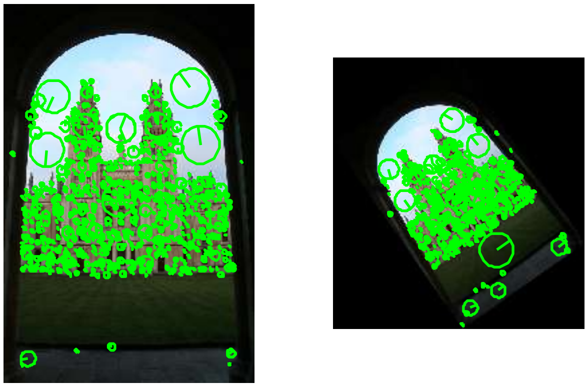

Fig 4.2: SIFT Feature detection with synthetic pair

In the proceeding stage, now we will display the pair of real image and reference template which are o be compared both, as shown in Fig. 4.3.

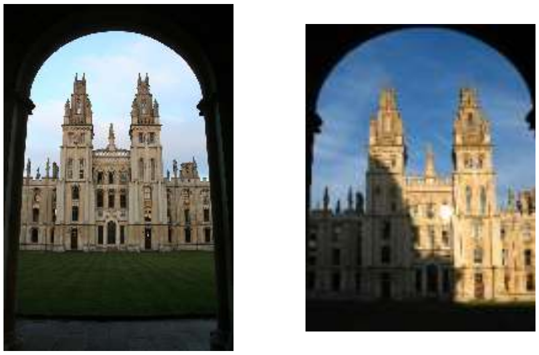

Fig 4.3: Original Images Real pair

In the next stage once again will generate SIFT features for both reference template and the original input image by applying the algorithm, as shown in Fig. 4.4.

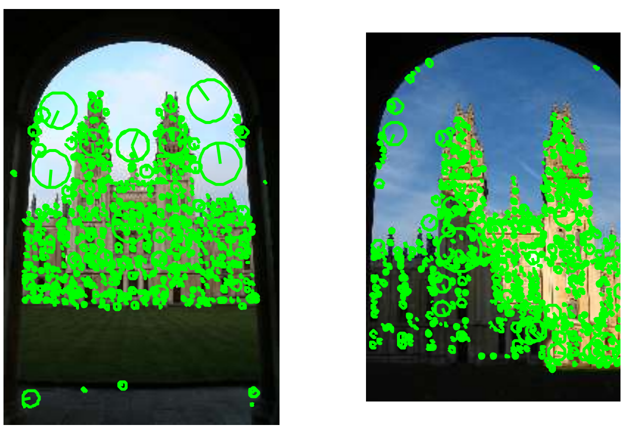

Fig 4.4: SIFT Feature detection of Real pair

The Fig. 4.5 represnts the SIFT features points for the reference image, which applies existing SIFT algorithm.

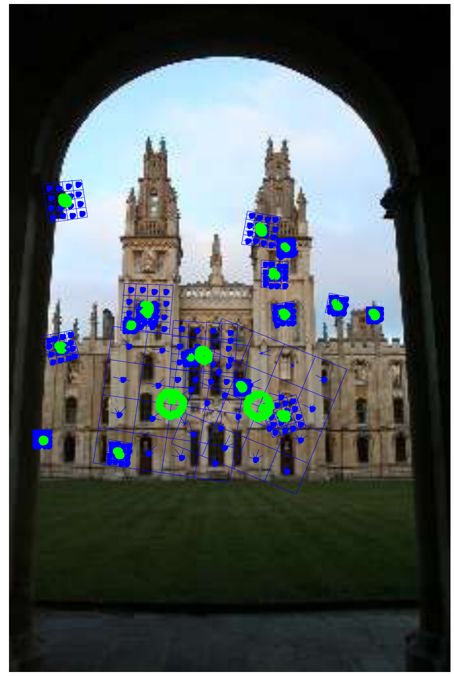

Fig 4.5: Existing SIFT Feature detection

In the next figure it depicts feature matching comparison with Nearest Neighbor test, which compares the keypoints of a reference image with original input image. Fig. 4.6 shows the output for this comparison.

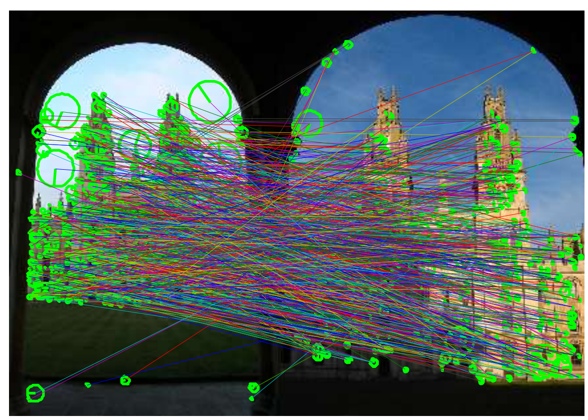

Fig 4.6: SIFT Descriptor Matching with Nearest neighbor test

In the next stage we will compare the feature keypoints by considering second nearest neighbor test, as shown inf Fig. 4.7, it discards irrelevant matching, so that it can reduce computation.

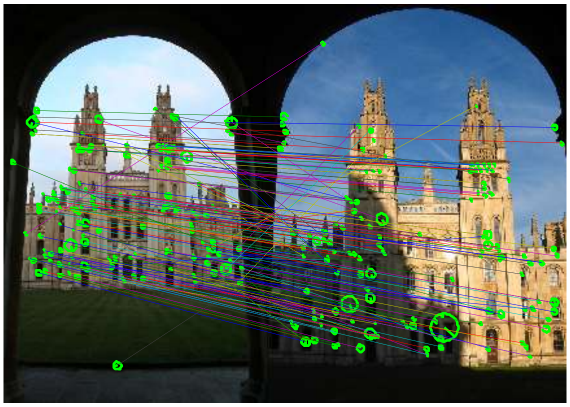

Fig 4.7: SIFT Descriptor Matching with Second Nearest neighbor (Lowe's) test

In the last stage of comparison, it verifies whether the matched points are correctly matching the corresponding feature keypoints, by performing the Geometric verification as shown in Fig. 4.8

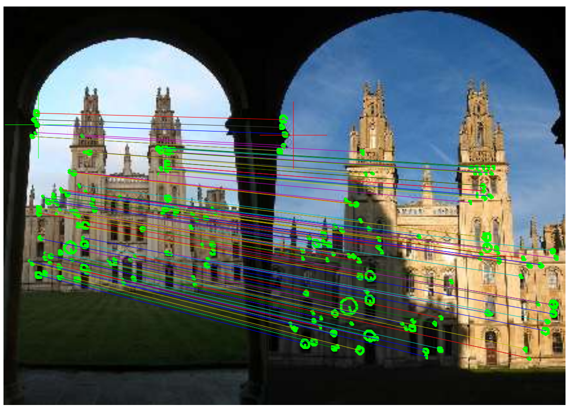

Fig 4.8: SIFT Descriptor Matching with Geometric verification

Performance Analysis:

The proposed project mainly works on Modified SFIT algorithm, which is the advanced version of exisiting SIFT. The SIFT algortihm performs extraction of keyfeatures and finds descriptors for the corresponding features, which are invariant to intensity changes, enlarging and revolutions. But SIFT alogrithm generally does not work well lighiting changes and blur. So to overcome this drawback we have choosen Modified SIFT wich gives better result than the earlier one interms of accuracy.

In the following sections we have taken some of the input images from the database which has been executed and we also noted down the corresponding readings which are listed as below.

In table 4.1 we have taken three different set of images, and hence we have obtained corresponding outputs. The graphical representation for this table will be shown in fig.4.9; here it shows variations of feature points with respect to each of the algorithm and also the features obtained by the Modified SIFT algorithm are more than the earlier method.

Table 4.1: Total number of feature comparison

Sl. No.	Image Size	Total Features	SIFT feature	Modified-SIFT Feature
1	768 x 1024 (Set_1)	2872	1731	2767
2	1024 x 767 (Set_2)	2215	1441	1760
3	1024 x 768 (Set_3)	2873	1734	2827

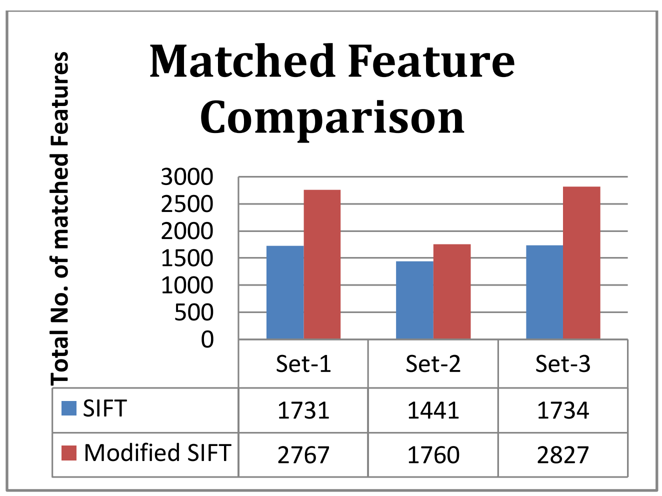

Fig. 4.9: Bar graph for the Feature comparison of SIFT's and Modified SIFT.

As listed in the table 4.2 gives a summary for the process time taken by the each method to perform matching algorithm and its graphical representation is also shown in fig. 4.10. From this graph we can easily conclude that the time taken by proposed method is less than the earlier one, as it only concentrates on the maximum peak threshold features. Thus reducing the computational overhead.

Sl. No.	Image Size	Require time to match	
		SIFT	Modified-SIFT
1	768 x 1024 (Set_1)	8.22 ms	6.61 ms
2	1024 x 767 (Set_2)	7.44 ms	6.12 ms
3	1024 x 768 (Set_3)	8.55 ms	6.99 ms

Table 4.2. Process time comparison

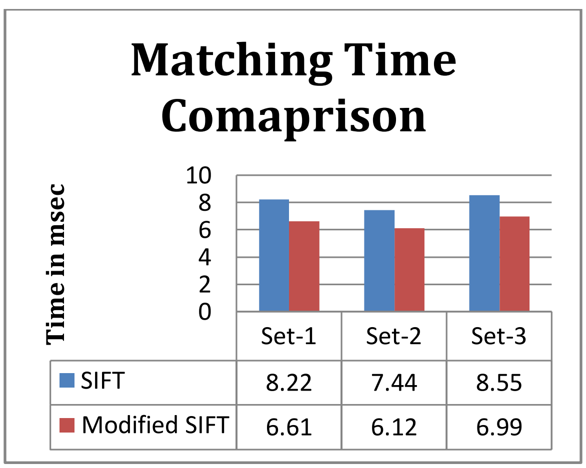

Fig. 4.10: Bar graph for the Matching time comparison of SIFT's and Modified SIFT.

The table 4.3 gives a comparison of accuracy for both SIFT and Modified SIFT algorithm. Here it compares accuracy based upon the number of features matched between the input image and reference template.

Sl. No.	Image Size	SIFT Accuracy %	Modified-SIFT Accuracy %
1	768 x 1024 (Set_1)	36.42%	99%
2	1024 x 767 (Set_2)	51.18%	98%
3	1024 x 768 (Set_3)	35.92%	99%

Table 4.3. Accuracy comparison

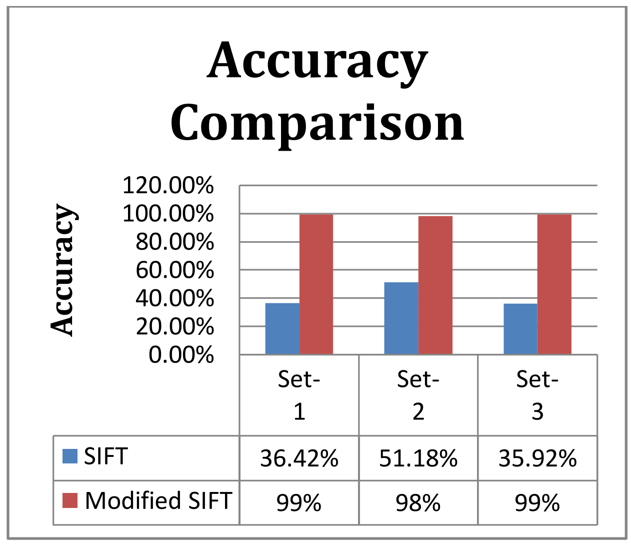

Fig. 4.11: Bar graph for the Accuracy comparison of SIFT's and Modified SIFT.

Based on the matched features we have calculated the accuracy and its graphical representation is shown in fig. 4.11, it depicts that accuracy calculated for each set will be greater than that of SIFT method.

V. Results and Discussion

Conclusion:

The proposed project mainly match respective image taken by the robot camera with Modified SIFT features of an image. It will be very beneficial to detect the accurate location of the image object by using Modified SIFT algorithm. This algorithm helped me a lot to perceive the various conditions of image recognition. Thus, this algorithm provides an efficient way to extract the local features from an image.

Future Scope of the work:

The proposed system we are concentrating on only Local features of the image, hence in future this algorithm performance can be enhanced by working with the Global features. These Global features are constructed by considering the whole image.

REFERENCES

[1] T. Ito, K. Yamada, and K. Nishioka, "Understanding driving situationusing a network model," Intelligent Vehicle, pp 48-53, 1995

[2] N. Matthews, P. An, D. Charnley, and C. Harris, "Vehicle detectionand recognition in greyscale imagery," Control Engineering Practice, Vol. 4, pp. 473-479, 1996

[3] S. D. Buluswar and B. A. Draper, "Color machine vision forautonomous vehicles," J. System Architecture, pp. 317-325, 1997.

[4] M. Bertozzi, A. Broggi, and S. Castelluccio, "A real-time orientedsystem for vehicle detection," J. System Architecture, pp. 317-325, 1997

[5] D. G. Lowe, "Object recognition from local scale-invariant features,"International Conference on Computer Vision, Corfu, Greece, pp.1150-1157, September 1999.

[6] W. V. Seelen, C. Curio, J. Gayko, U. handmann, and T. Kalinke, "Scene analysis organization of behavior in driver assistance system," Proceedings of IEEE International Conference on Image Processing, pp 524-527, 2000

[7] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W.Y.Seelen, "An image processing system for driver assistance," Imageand Vision Computing, Vol.18, pp. 367-376, 2000.

[8] C. Schmid, R. Mohr., and C. Bauckhage, "Evaluation of interest pointdetectors," International Journal of Computer Vision, Vol. 37, No. 2,pp. 151–172, 2000.

[9] K. Mikolajczyk and C. Schmid, "An affine invariant interest pointdetector," European Conference on Computer Vision, Copenhagen, Denmark, pp. 128-142, 2002.

[10] D. G. Lowe, "Distinctive image features from scale-invariantkeypoints, International Journal of Computer Vision," Vol.60, No. 2, pp. 91-110, 2004.

[11] Haoxiang Lang, Ying Wang and Clarence W. de Silva, "Vision Based Object Identification and Tracking for Mobile RobotVisual Servo Control" 2010 8th IEEE International Conference on Control and Automation Xiamen, China, June 9-11, 2010

[12] JaspreetKaur and NavjotKaur, "Modified SIFT algorithm for Appearance based Recognition of American sign language", IJCSET Vol 2, Issue 5.1197-1202, May 2012

[13] Mohd. Ahmed, S. A. Patekar, and MandarSohani, "Short Listing Likely images using proposed Modified SIFT together with conventional SIFT for quick Sketch to photo matching", International Journal in Foundations of Computer Science & Technology (IJFCST), Vol. 3, No.4, July 2013.