# Scale-Invariant Object Identification for Jointly Addressing Issues in Moving Oject and Background

Dr. Pushpa D.
*Asst. Prof*: MIT, Mysore, India
E-Mail: pushpad046@gmail.com

Dr. H.S. Sheshadri
*Prof*.: Dept. of Electronics & Communication Engg.
PES College of Engineering
Mandya, India
E-Mail: hssheshadri@gmail.com

***Abstract*— the area of video process pertaining to analysis of the real-time object behaviour is increasing in its pace. However, such process is also accompanied by various critical challenges e.g. variation in illumination condition, complete or partial occlusion, and uncertain mobility of the object. Although, in the last decades there has been archives of research papers discussing about the object detection, but there are only few literatures pertaining to the precise moving object detection with moving background. Therefore, the core objective of this paper is to present a technique that efficiently performs moving object detection for the moving background using enhanced version of Scale-Invariant Feature Transform. The outcome of the study was compared with the most significant work to find it outperforms the existing system with respect to processing time and thereby it is quite found suitable for real-time processing of object detection.**

***Keywords-Component; Image Processing, Moving Object Detection, Object Tracking, Occlusion, Video Surveillance***

## I. INTRODUCTION

The evolution of Digital Image Processing has made a significant contribution in the area of computer vision and its related applications [1]. Computer Vision has also make a significant progress in identifying specific objects from the given scene information. The prime purpose of adopting object detection is to identify the specific object on the most complex background scene. The process is however not that easy as there are various factors that affect the outcome of object detection. There are various military applications where video scene or image is captured from unmanned vehicle [2] and the captured archives of the scene are subjected for instantaneous object detection. The complexity in the object detection is basically the acquisition technique as there are various possibilities of object information acquisition [3] e.g.

- stationary object with static background
- moving object on static background
- static object with moving background
- Moving object with moving background.

As discussed in prior studies that there are various possibilities of object mobility and their orientations e.g. i) stationary object and stationary background, ii) moving object with stationary background, iii) stationary object with moving background, and iv) moving object with moving background. Although to some extent, extracting the foreground from background can be carried out on i) stationary object and stationary background, ii) moving object with stationary background, and iii) stationary object with moving background. But in the 4th case, i.e. moving object with moving background is quite challenging scenario where the normal segmentation process or geometrical homography based process will not yield much productive outcomes. The prior studies has introduced various techniques like e.g. i) homogeneous object detection using low-level features and corner detection techniques [8], ii) heterogeneous object detections using unscented Kalman filter[9], iii) detection of single object from crowd from multiple views using homography-based approach [10]. However, the approaches mentioned above works suitably fine with i) stationary object and stationary background, ii) moving object with stationary background, and iii) stationary object with moving background only, but not on challenging cases like moving object with moving background.

The area of video surveillance system specifically pertaining to object detection encounters several of research issues which are critical to be satisfied from the viewpoint of the outcome analysis. Followings are the brief discussion of some set of problems that has been identified and is subjected to be addressed in the proposed study:

- **Variation in Illumination Condition**: Usually the security cameras used for the purpose of video surveillance system are not high end camera. Hence, such low resolution cameras are quite sensitive to the lighting condition of the surveilled scene. In case of poor illumination condition, the object detection usually fails, resulting in generation of false alarms of object detections. Generally, the research community considers the RGB frames and performs processing it by converting it to grey scale. Hence, poor illumination condition results in generation of less pixel information by missing out some significant feature information.
- **Occlusion**: Occlusion usually occurs when the scene is heavily crowded and a feed captured by the camera is not able to identify the occluded or partially occluded objects. Although the security cameras are mounted in top position of the scene, but in case of occlusion, the camera is not able to detection the object moving behind another object resulting in complete occlusion. In this case, the detection

of the underlying object completely fails. In case of partial occlusion, the complete criteria to explicitly define the object fails and results in false positive.

- **Detection Method**: At present, majority of the existing research focuses on detection all objects of similar types (i.e. homogeneous objects). The problems becomes must worst, when multiple different moving objects are considered for case study (heterogeneous objects). In case of low-lightening condition, occlusion, understanding the heterogeneity of the object is computationally challenging task in object detection. Majority of the existing research attempts has focused on datasets with two-three dimensional video feeds, but in reality, it is found that similar event is captured from multiple cameras thereby increasing the dimensionality. Hence, detection a single object from multiple objects from multiple video feeds is another computationally challenging process to accomplish.

Hence, it is clearly seen from the above mentioned points that there is a need of a system that can overcome all the above mentioned challenges or research issues. The problem statement of the study can be thereby re-formulated as "*It is computationally challenging task to detect a single object most precisely from multi-dimensional feeds in presence of occlusion (full or partial, low lightening condition and different mobility patterns of the background and objects*." This paper presents a novel technique that can perform detection of an object considering moving object and moving background. Section II discusses about existing literatures followed by discussion of assumptions and issues in Section III. Section IV discusses about proposed model followed by discussion of an implementation techniques in Section V. The result discussion is done in Section VI, while summary of the paper is done in Section VII.

## II. Related Work

This section will briefly discuss about the relevant research work being carried out in recent past. Mori et al. [11] have presented a technique for detection of moving objects followed by classification based on grid trajectories. The authors have designed a hardware based model and have evaluated with respect to recall rate, precision rate, and f-measure. Zhao et al. [12] have also carried out the object detection using the concept of background subtraction using learning based method. The core modelling was done based on sparse matrix representation. The outcome of the study shows much lesser false positive. Wu et al. [13] have also worked on moving object detection using the shadow elimination process followed by reconstructions of prospect. The outcome of the study was evaluated with respect to the visual outcomes where the identified moving objects were shown with bounding boxes. Work on similar direction was carried out by Zhou et al. [14], where the authors have presented a technique to identify the outliers. The study has also introduced a segmentation process using minimal rank modelling scheme for identifying the moving objects from the complex background process. However, the technique was found not to be quite compatible for real-time moving object detection.

Yi et al. [15] have designed a technique that can perform identification of a moving object from the real-time video captured from the trusted handheld devices. The authors have used Single Gaussian Model, C++, and OpenCV library. Xia et al. [16] have presented a work that identifies the foreground that moves using both distance and time-based approach. Just like Yi et al. [15], this work has also used Gaussian Mixture Model. The study has adopted the usage of time based approach for saliency map for performing significant background information. The outcome of the study was evaluated with respect to ROC curve. Thorat and Nagmode [17] have used statistical technique to perform moving object detection as well as tracking. The authors have used correlational-based approach over the video frames to perform moving object detection. Kryjak and Gorgon [18] have used FPGA for moving object detection. Luo [19] have modelled the background for identifying the moving object using statistical analysis. Patel et al. [20] have used a unique approach of averaging filter as well as Gaussian filter for identifying the moving background that has slight variational information. The next section discusses about the assumptions and the issues being identified in the proposed study followed by the discussion of the proposed system that mitigates the issues

## III. Assumptions and Problems

The assumptions of the proposed system are as follows:

- The centroid intensity is assumes that intensity of the corner is just an offset from the center of gravity of the frame.
- The frames extracted from the video are free from any noise.
- The frames extracted are not subjected to any preprocessing operations.

The problems that these studies of the thesis are addressing are related to the mobility of an object as well as background. Following are the issues that occur owing to this:

- The moment an object as well as background starts moving together, it becomes very difficult for the conventional video segmentation as well as background subtraction is not directly applicable.
- It is quite a difficult task for the system to perform background subtraction specifically for the object that is in constant move.
- The situation turns worst when the object as well as background are in synchronal mobility with each other, which makes the system complicated to perform object detection
- Toggling from foreground to background subtraction process is good option for static background but is time consuming for mobile background objects.
- It is quite a difficult task to ensure faster processing of the algorithm that can perform detection of moving objects.

## IV. PROPOSED SYSTEM

The core objective of the proposed system is to develop a technique that can perform object detection in a special case of moving object with moving background. Detection of moving object itself is a challenging task and more issues are added when the background is also considered to be moving. The proposed system is designed based on the recent protocol developed by Ethan Rublee et al. [21] in 2011 who have discussed about the new technique called as Oriented FAST and rotated BRIEF (Binary Robust Independent Elementary Features). The prime purpose of the proposed model is to enhance the conventional FAST technique and enhancing the computation to superior level for oriented BRIEF feature. The idea of the proposed system is based on the SIFT (Scale-Invariant Feature Transform) detector, which is basically used as a visual descriptor. The proposed study also adopts the principle of affine homography concept for overcoming the problems associated with object detection when both object and its background are in mobile conditions. However, the proposed system in this part of the study chooses not to use conventional SIFT protocol as it is usually found to be computationally complex process. Hence, this has lead to the evolution of SURF protocol that has considerable computational speed compared to conventional SIFT. The prime motivation of adoption of substitution of SIFT protocols to something novel as if the case study of moving object with moving background is selected, it will also lead to considerable amount of inclusion of image noise.

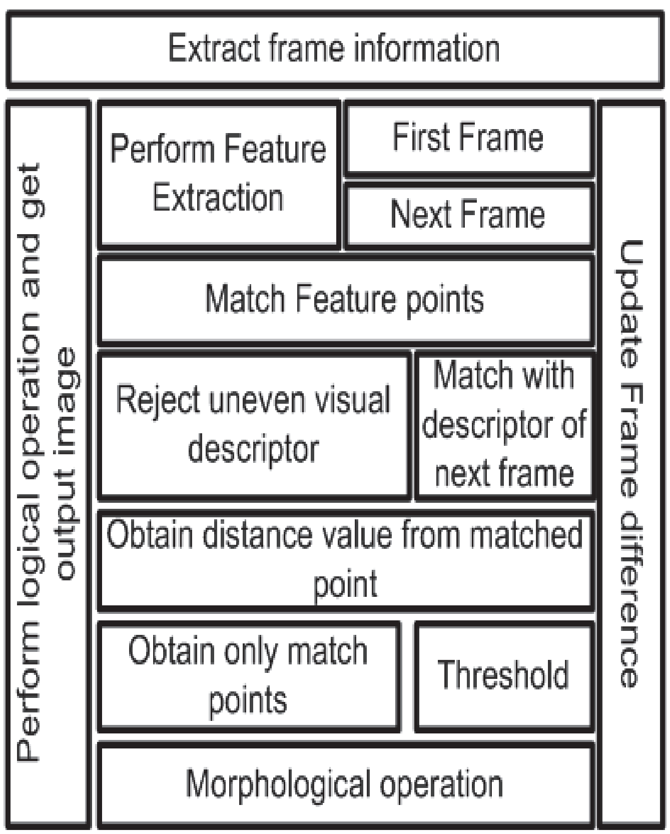


Figure 1. Adopted Architecture for Extensive Analysis

Fig.1 showcases the proposed architecture where the system considers the critical mathematical components of affine homography based transformation for carrying out tracking process of the moving object in each consecutive frame. The primary motivation for the above discussed architecture was also drawn from the study conducted by Deniz et al. [22], an author who has presented an algorithm for accurate universal mobility identification. The study also adopts considering rotated BRIEF and oriented FAST, which are certain sets of frequently adopted visual descriptors for selecting local and global feature. The design principles ensures that proposed system gives better detection and tracking for the moving objects even compared to the conventional SIFT method. It also addresses the issues of more rapidly occurring motion compensation. The entire operation is carried out on following steps e.g. i) attribute selection and computation, ii) performing affine homography transformation, and iii) performing detection of moving object on moving background.

The processing of the proposed system initiates after the features are extracted using proposed system. The feature extractions are done using visual descriptors between the two sequences of the frames. This operation is followed by comparing the unmatched features with the matched features using the visual descriptors. The methodology completely discourages getting the unmatched feature pairs as it will even lead to false positives and lead to wrong object detection. The feature pair where the positive match is found is further considered for the processing using visual descriptors as well as affine homography-based transformation. The next phase of the processing will include the entire extracted frame to undergo mathematical transformation process for the consecutive frames. This processing step significantly compensates the issues of dynamic motions that are involved in the datasets. The error rate is evaluated from the consecutive frames for examination of the potential information about the background objects.

Finally the background objects are eliminated for ensure precise foreground object detection as well as tracking. In this entire processing to evolve up with foreground and background objects from the dataset, it is very much feasible that specific levels of residues are remaining in the matrices because of the frame rate errors. Such remnant residues are checked and eliminated from the internal matrices using various morphological operations as well as logical operations in order to accomplish better precision rate. Hence, the proposed system offers better capability to the framework to generate precision moving object detection in moving background in presence of heterogeneous objects. The sole purpose of the proposed system is to evolve up with another new technique that can be adopted for investigation of the phenomenon of object detection under complex illumination condition as well as complex mobility scenario. The motive for this last phase of the study is to investigate the foreground as well as background object along with then associated complexities in terms of mobility.

## V. IMPLEMENTATION

The design principles of the proposed system discussed in the previous paragraph illustrates about the technique of moving object detection where the algorithm implementation is discussed in this section. The design of the proposed study was done using Matlab on conventional windows based platform. The system attempts to finds the local features in a given frame with an efficient performance as well as reduced computational cost. The study conducted in [23] was considered as a backbone to design the algorithm for evaluating the local feature points in the consecutive frames in highly precise manner. The methodology of the study [23] was similar to ring topology of 16 pixels that is located in corner. The system uses visual descriptors that understands the implication of corner and highlights it as a corner point if there is a group of around 12 redundant pixels in the all 16 locations of the ring topology. The brightest pixels are estimated and compared with other intensities of the pixels $I_c$ and the threshold $t$ and is represented as $I_c + t$, while the reverse case with darker pixels is represented as $I_c$-t. Hence, it is quite obvious that the proposed system mimics the design principle of Rosten Drummond technique discussed in 3rd study of ours. However, in this part of the extensive analysis, the study was carried out for 4 directions only. Therefore the global feature points can be represented with moments of the features as,

$$\phi = a\tan 2(\sum_{x,y} y.I(x,y), \sum_{x,y} x.I(x,y)) \quad (1)$$

In the above eq. (1), I(x, y) represents image intensity corresponding to a particular position of an object and is considered as a global feature points. The concept of feature points in this part of the study is very significant as it is used for visualizing the potential similarity of any particular object located in two consecutive frames with minor or major displacements. The process of the algorithm than performs extraction of the targeted features of an object which is actually a type of visual descriptor using binary strings and hamming distances. The proposed technique adopted has various advantages characteristics. It lowers down the resource consumption while performing the feature extraction process and hence the proposed system is considerably faster compared to any existing techniques of feature extraction from targeted objects. Not only this, the algorithm also perform spontaneous matching of the similarity point between any two consecutive frames for confirming that the two objects in two different frames is a same / different objects. The anticipated performances of the proposed algorithm are quite faster compared to the conventional SURF or SIFT techniques.

## VI. RESULT DISCUSSION

For the purpose of the performance comparative analysis, the outcome of the proposed system is compared with that of Yi et al. [15]. Yi et al. [15] have also worked on the similar problem considering moving objects detection. The author have performed modelling of their approach using double mode distribution of Gaussian approach considering temporal factor in order to resist the mixing of pixels related to the background and foreground objects.

Figure 2. Identifying moving objects on moving background

Fig.2 shows the visual outcomes of the proposed study. While checking the outcomes of the proposed system, the dataset was screened for various problematic illumination scenarios, partial, full occlusion, and mobility of multiple moving objects. The result shows the sequences of the $i^{th}$ frame, $(i+1)^{th}$ frame, $(i+2)^{th}$ frame, and $(i+3)^{th}$ frame. The outcome of the proposed system was evaluated with respect to the accuracy in performing moving object detection on moving backgrounds. The multimedia clip used as a dataset for the proposed experiment consists of 5 moving objects and therefore, the evaluation considers formulations of performance parameters that will assist in performing benchmarking.

Following are the performance parameters:

- Originally detected region of an object=$\Delta_1$
- Falsely detected region of an object = $\Delta_2$
- Missing Object= $\Delta_3$

This technique allows the actual region of an object ($\Delta_4$) manually for estimating the rate of object detection ($\Delta_5$),

$$\Delta_5 = \Delta_1 / \Delta_4 \quad (2)$$

Moreover the error occurred owing to false rate of detection of an object can be represented as $\Delta_6$,

$$\Delta_6 = \Delta_2 / (\Delta_1 + \Delta_2) \quad (3)$$

The evaluation can also study the features that are not the object and hence its rate of detection can be represented as $\Delta_7$

$$\Delta_7 = \Delta_3 / \Delta_1 \quad (4)$$

All the above performance parameters are used for the purpose of understanding the efficiency of the proposed extended analysis on moving object on moving background. The analysis is carried out for 10 iterations to arrive to following outcomes.

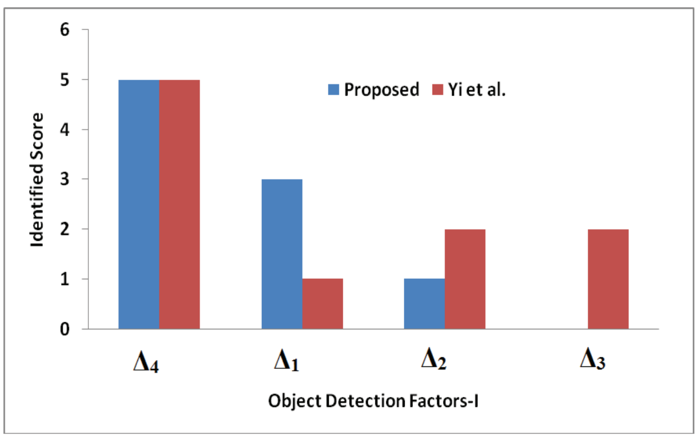


Figure 3. Performance Analysis-I

Fig.3 discusses about the effectiveness of the proposed system that is found benchmarked with study of Yi et al. [15]. The actual test scenario consists of 5 objects from which 3 objects are able to be identified in the outcome analysis. The false identification was shown for only one object in the test scenario. However, the authors Yi et al. [15] have adopted the conventional techniques of motion compensation by using Gaussian distribution. This is one of the reasons, why the work done by Yi et al. [15] seems to underperform as compared to the proposed system. Therefore, by using Oriented FAST and rotated BRIEF, the system performance drastically increases.

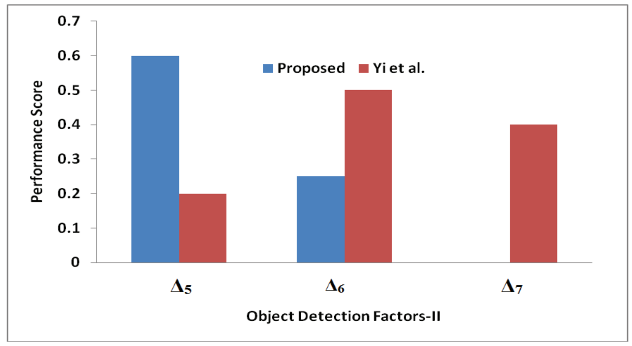


Figure 4. Performance Analysis-II

Another set of outcome is exhibited in Fig. 4 that shows the individual performance of proposed system exhibits the comparative performance analysis, where it can be seen that proposed system shows superior detection capability with considerable less error in object identification process. False positives outcomes for the proposed system are found to be zero.

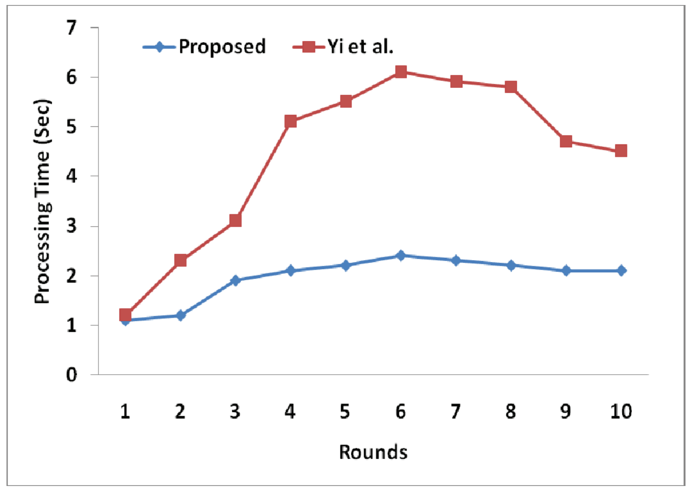


Figure 5. Evaluation of Processing Time

The outcome of the processing time for both the techniques is exhibited in Fig. 5. Owing to inclusion of Oriented FAST and rotated BRIEF techniques, the proposed system performs faster computation as compared to the existing system. The cumulative time for processing the algorithm for proposed system has almost linear behaviour.

## VII. CONCLUSION

This study has presented an extended analysis of the proposed study where the homography based single object detection and tracking is enhanced for moving object with moving background case study. The proposed system has used Oriented FAST and rotated BRIEF to find that the technique is efficiently faster and cost effective. The comparative analysis also proved that proposed system bears better detection performance and efficient processing time using affine homography based transformation technique. At present our targeted object to be detected is bigger in size. Our future work will be focused to include detection of moving object of much smaller in size and will also include heterogeneous object.

## REFERENCES


[1] A. Das, Guide to Signals and Patterns in Image Processing: Foundations, Methods and Applications, Springer, 2015

[2] M. LaFay, Drones For Dummies, John Wiley & Sons, 2015

[3] S. H. Shaikh, K. Saeed, Nabendu Chaki, Moving Object Detection Using Background Subtraction, Springer, 2014

[4] T. Bouwmans, F. Porikli, B. Hoferlin, A. Vacavant, Background Modeling and Foreground Detection for Video Surveillance, CRC Press, 2014

[5] M. Zhi, X. Zhang, “Object Detection from replay shot in tennis video”, Taylor and Francis-Industrial Engineering and manufacturing Technology, 2015

[6] R. Kraemer, M. Katz, Short-Range Wireless Communications: Emerging Technologies and Applications, John Wiley & Sons, 2009

[7] Teutsch, Michael, Moving Object Detection and Segmentation for Remote Aerial Video Surveillance, KIT Scientific Publishing, 2015

[8] D. Pushpa, H.S. Sheshadri, “Precise multiple object identification and tracking using efficient visual attributes in dense crowded scene with regions of rational movement”, *International Journal of Computer Science Issues*, Vol. 9, Issue 2, No 2, March 2012

[9] D. Pushpa, H.S Sheshadri, “Semantic Analysis of Precise Detection Rate in Multi-Object Mobility on Natural Scene using Kalman Filter”, *Springer-Emerging Research in Electronics, Computer Science and Technology Lecture Notes in Electrical Engineering*, Volume 248, 2014, pp 217-226

[10] D. Pushpa, H.S. Sheshadri, “Homography based Multiple Moving Objects Using Multi-Dimensional Event Feeds, International Journal of Computer Science”, *Information Technology & Security*, Vol.3 No. 5 October 2013

[11] T. Mori, T. Sato, H. Noguchi, “Moving Objects Detection and Classification Based on Trajectories of LRF Scan Data on a Grid Map”, *IEEE- International Conference on Intelligent Robots and Systems*, pp.2606-2611, 2010

[12] C. Zhao, X. Wang, andW-K Cham, “Background Subtraction via Robust Dictionary Learning”, *EURASIP Journal on Image and Video Processing*, 2011

[13] W. Wu, J. Shao, and W. Guo, “Moving-object Detection Based on Shadow Removal and Prospect Reconstruction”, *International Conference on Artificial Intelligence and Soft Computing*, vol.12, 2012

[14] X. Zhou, C. Yang and W. Yu, “Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation”, *arXiv*, 2012

[15] K. M. Yi, K. Yun, S. W. Kim, “Detection of Moving Objects with Non-Stationary Cameras in 5.8ms: Bringing Motion Detection to your Mobile Device”, *IEEE-Open Access on Computer Vision Foundation*, 2013

[16] Y. Xia, R. Hu, Z. Wang, and T. Lu, “Moving Foreground Detection Based On Spatio-temporal Saliency”, *International Journal of Computer Science*, Vol. 10, Issue 1, No 3, January 2013

[17] S. Thorat, M. Nagmode, “Detection & Tracking Of Moving Object”, *International Journal of Innovative Research in Advanced Engineering*, Vol.1, Iss.1, April 2014

[18] T. Kryjak_, M. Gorgon, “Real-Time Implementation Of Moving Object Detection In Video Surveillance Systems Using FPGA”, *IEEE-Computer Science*, Vol. 12, 2011

[19] Y. Luo, “Moving Object Detection based on Background Modeling”, Institutionen för informationsteknologi -Department of Information Technology,2014

[20] C. I. Patel, S. Garg, T. Zaveri, and A. Banerjee, “Top-Down and Bottom-Up Cues Based Moving Object Detection for Varied Background Video Sequences”, Hindawi Publishing Corporation, Advances in Multimedia, 2014

[21] E.Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: An efficient alternative to SIFT or SURF", IEEE International Conference on Computer Vision, pp. 2564-2571, 2011

[22] O.Déniz, G. Bueno, E. Bermejo, and Rahul Sukthankar, "Fast and accurate global motion compensation”, Pattern Recognition, Vol. 44, No. 12, pp. 2887-2901, 2011.

[23] E.Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection", Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol.32, No. 1, pp. 105-119, 2010