A survey about object detection and classification in image processing

[1] S.Shelgin, [2] K.Shanmugam, [3] Dr.B. Vanathi
[1] PG Scholar, [2] Assistant Professor, [3] Professor & Head
[1,2,3] Computer Science and Engineering,
[1,2,3] Valliammai Engineering College, Chennai, India

***Abstract*— In this paper the various object detection and classification algorithms are analyzed which are used to detect an object using image processing. The images used for processing were taken from the surveillance camera placed in areas such as traffic analysis, border area, ATM centers, etc. Detection of moving object is a challenging task, for this purpose the object detection and classification of various images are performed. The object detection is done for checking whether there is any object in the area or not and also to locate the position of the object. The various detection algorithms are analyzed with their performance. The object is also classified into various categories such as human, animals, trees, etc. For this we have analyzed the various object classification algorithms and the performance is obtained. Further the images taken from the outdoors, due to some environmental factor get degraded. This can be rectified using dehazing algorithms and in this paper the various dehazing algorithms are compared and analyzed with their performance.**

***Index Terms*— Object detection; object classification; Haze removal.**

I. Introduction (Heading 1)

In image processing, the images are processed using some mathematical functions whose input are the image itself, which is either obtained from the surveillance video or a photograph. The image processing is usually referred as digital image processing since the image obtained is converted into digital format. The image in the digital format is represented as the 0's and 1's called binary code [13]. The sequences of 0s and 1s that constitute information are called bytes. In this paper, the object detection and object classification are performed with the sequence of images which are generally known as videos. Each sequence in the image is known as a frame. Each frame that is converted into digital format for processing the image is called as digitization process [12]. Digitizing or digitization is the representation of an object, image, sound, document or a signal (usually an analog signal) by a discrete set of its points or samples. An image obtained from a surveillance camera consists of two parts and they are background and foreground image. The ultimate aim of object detection is to find out the foreground image and locate their position, size and so on. In order to locate the foreground image the background elimination is to be performed which is known as background subtraction [9][[10][15][20]. Generally the foreground objects are the moving objects such as peoples, animals, birds, vehicles, etc. There are two stages of object detection in image processing that can be finally used for object tracking or some other operations like object recognition, image quality assessment, and so on[1][7][25]. The image that is affected due to environmental factors, such as fog, mist, rain, etc., is removed by using the dehazing algorithms. The images gets degraded due to the irradiance of the light [13][14][22].The dehazing process is used in many computer vision and image editing process[27][28].

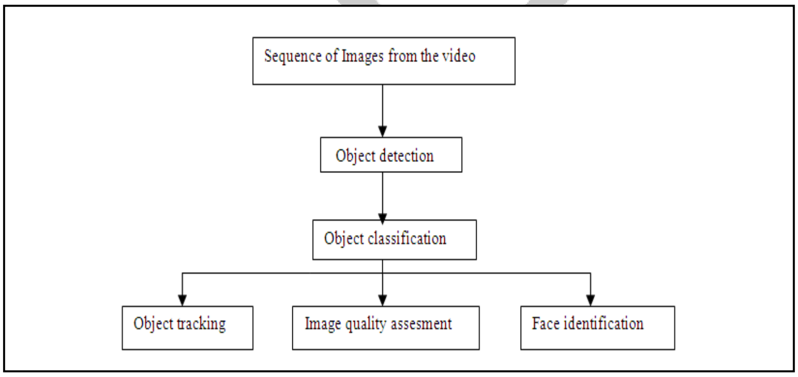

Fig. 1. Basic Image Processing Process

In this paper, we just described about the various algorithms used for the following two techniques.

- Object detection.
- Object classification.
- Dehazing Algorithms.

Object Detection

The object detection is the computer technology that is related to computer vision and image processing that deals with the detection of the instance of the semantic objects of certain classes such as human, car, buildings, etc., in digital images and videos [14]. The object detection includes the face detection, pedestrian detection and so on. The object detection is very useful in the field of image retrieval, video surveillance applications.

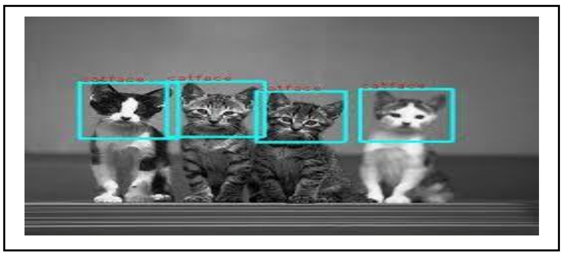

Fig. 2. Object Detection Process

In this paper the following algorithms are analyzed.

- Diffusion based salient object detection
- Optical flow
- Contour based detection
- Affine scale invariant transformation

Object classification

The object classification is the process of classifying the image such as vehicles, human, animals, and any other moving objects [4]. The object classification is done to find the object accurately from the image. The image classification is also known as pattern recognition and the pattern refers to the object in the image [12]. The classification is further divided into supervised classification and unsupervised classification [5-8]. In supervised classification, the classification is done by using a statistical character of the image, which is known as a signature. Using this signature and the analyst-specified training data the object can be classified. In unsupervised classification, the object is classified using the samples of a large number of unknown pixels in the image and then divided into a number of classes. The object classification is very useful in computer vision application.

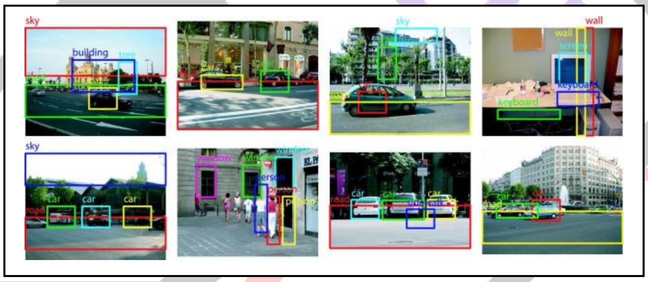

Fig. 3. Object Classification Process

In this paper the following algorithms are analyzed.

- Feature based classification
- Shape based classification
- Texture based classification
- Color based classification

Dehazing Algorithm

The image, taken in the outdoor is considered to be degraded if it contains the environmental factors such as fog, mist, etc. This degradation will affect the quality of the image[16][27]. Hence the factors that affect the originality of the image should be removed by using the dehazing algorithms. The following are the dehazing algorithms.

- Color Attenuation Algorithm
- Dark Channel prior algorithm

II. OBJECT DETECTION METHODS

Object detection is the process of locating the region of interest called the object in an image. The following is the detailed explanation of various algorithms used in object detection.

Diffusion Based Salient Object Detection

In [1], diffusion based salient object detection algorithm is used for learning the optimal seeds of the images in order to locate the human eye fixations. This method uses the bottom up saliency detection technique that uses local measures of center-surround contrast or global measures of stimulus uniqueness for predicting the eye fixation and this method is useful for high level vision task such as object recognition and localization. An image is partitioned into superpixels (primitives of the image such as shape and segmentation) or image patches. They also include spatial propagation of saliency information in order to overcome the propensity of bottom-up saliency to respond more to the edges than object interiors.The spatial propagation uses the object features. The superpixels are mapped into a graph. Here the superpixels act as the node and edge strengths proportional to superpixels. The saliency information is allowed to propagate over the graph until it reaches the equilibrium state and this mapping is called an object saliency map. This method is used to learn the optimal seeds for object saliency by combining pre-attentive saliency maps and mid-level vision cues for object perception. Every superpixel has two types of features that have to be computed. They are the bottom up saliency of superpixel region and a set of midlevel vision feature information.

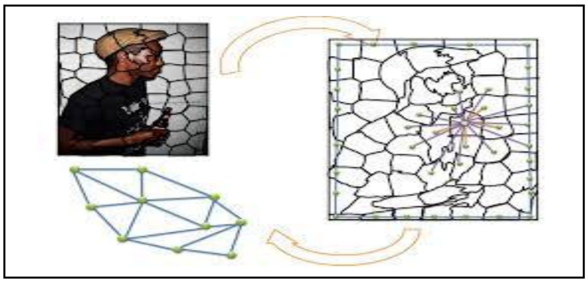

Fig. 4. Saliency Seeds, Propagation Graph And Object Saliency

The combination of these two features allows a best discriminates between the object and the background and this combination is known as large-margin formulation of the discriminant saliency principle as shown in fig 4. The image is represented as a graph over the set of superpixel and is efficiently used for detecting an object in the image. It is also found that the state-of-art of the result is maintained on a number of object saliency datasets.

Optical Flow Method

In [2], the optical flow method is used to detect the moving object with camouflaged color. Here the optical flow model is used to distinguish the motion pattern of the object and background of the object where the foreground and background objects have similar color in the image. The motion pattern of the image is done by using the corner detection and LK based optical flow method is used as it has less computational time. This is done by pyramid implementation. Thus the magnitude and the location of the image are obtained and it is clustered in to the motion pattern in order to detect the object. Finally the Kalman filter is used to improve the performance of the object detection. The kalman filter uses the state variable, which are the location and the scale of the object. In [26], the optical flow estimation is provided for the oversampled video by using the Lucas-Kanade algorithm. The estimation is accumulated and refined from high frame rate to standard frame rate. The optical flow method provides high accuracy in detection. The optical flow was easily influenced by noise impact and this is the major disadvantage in the optical flow method.

Contour Based Detection

In [3], the shape based contour is extracted. This method uses the border or the edges of the background images. This technique is based on the information provided by the object boundaries. Boundary features provide more precise shape information. First a rough contour is extracted using canny edge detection method by edge pixel values. After that parameters are estimated using shape matching. A shape is represented by a discrete set of points sampled from the contours on the shape. These points can be on internal or external contours. Tracking is finally completed by elastic shape matching for extracting the exact contour. This approach has advantage like rotation, biased occlusion handling, and translation. But in crowded outdoor environment more robust technique is required. Contour based tracking is useful as the shape of the object is detected.

Affine Scale Invariant Feature Transformation

In [4], the object is detected using the full boundary information of the image along with the affine scale invariant feature transform (ASIFT) and a region merging algorithm. The ASIFT uses six affine parameters such as two parameters for translation and camera axis orientation and one parameter for zooming and rotation of the camera. Extracting the keypoints from an image is very important and valuable for detecting the object in the image. The detection and reorganization of the object is done by using the keypoints and segmentation algorithm and the result of this method is very accurate. If the keypoints are obtained correctly then they provide the best information from the image. In this paper the authors have combined the results of ASIFT and a region merging algorithm in order to recognize the object and detect them with full boundary information. Here the regions need not to be specified by the user because of the usage of the keypoints from the result of ASIFT and this algorithm is an automatic algorithm. The ASIFT is the extension of SIFT. The keypoints are used for providing the matching between different images of an object. The ASIFT algorithm uses the affine camera model, scale space extremum detection, accurate keypoint localization, assigning an orientation and a keypoint descriptor. Finally the region merging algorithm that is done based on maximal similarity and is used for the detection of the object. ASIFT is more effective for detecting the object and can be more robust in the changes of images. This algorithm is more accurate in detecting as well as recognizing the object. The extremum of DOG across scales, the keypoints obtained after removing the low contrast points and the keypoints obtained after removing the edge response is represented in fig 5 (a)(b)(c) respectively.

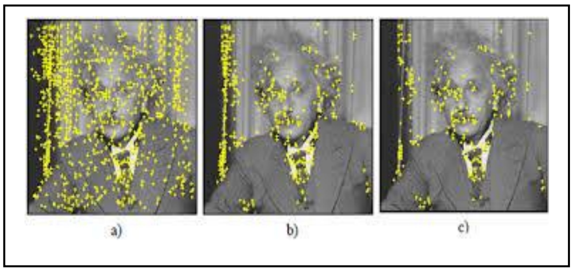

Fig. 5. (a) Extremum Of DOG Across Scales.
(b) Remaining Keypoints After Removal Of Low Contrast Points.
(c) Remaining Key Points After Removal Of Edge Responses.

TABLE I. COMPARISON BETWEEN DIFFERENT OBJECT DETECTION ALGORITHMS

Algorithm	**Advantage**	**Disadvantage**
Diffusion Based Salient Object Detection	➢ Efficiently used in the detection of an object. ➢ The stare-of-art of the result is maintained.	➢ High computational time.
Optical Flow	➢ Efficiently used in detecting the object that has camouflaged color.	➢ Easily affected by noise.
Contour Based Detection	➢ Efficiently used for detecting the moving object that has occlusion problem.	➢ Cannot be used in crowed area. ➢ Only shape of the object is detected.
Affine Scale Invariant Transformation	➢ Most accurate detection algorithm.	➢ Shape of the object cannot be detected.

III. OBJECT CLASSIFICATION ALGORITHMS

The object classification is the process of classifying the image such as vehicles, human, animals, and any other moving objects. The object classification is finding the object accurately from the image. The following is the detailed explanation of various algorithms used in object detection.

Feature based classification

In [5], the feature based classification of the object is performed for classifying the image and match the image with the required result. In feature based classification different features like skin color, height, weight, width, etc., of the object is extracted. Here the author describes the moving region segmentation that is implemented using background subtraction method since the object's background should be eliminated and for multicolor environment the GMM is used. The feature is extracted by finding the centroid of the object and the bounding box provides the spatial position and matching is done with the extracted feature. In [19], describes the classification of unique synthetic aperture radar image (SAR). The image is partitioned and the ROI is obtained and the segment is modified based on the features of the images and thus classification is done. Feature based classification is useful when the shape of the object is not clear. This algorithm is not suitable for different environment and weather condition.

Shape based classification

In [6], Different descriptions of shape information of motion regions such as representations of points, box and blob are available for classifying moving objects. Input features to the network is mixture of image-based and scene-based object parameters such as image blob area, apparent aspect ratio of blob bounding box and camera zoom. Classification is performed on each blob at every frame and results are kept in histogram. In [25], the classification is done for the retro-reflective traffic signs obtained from Moving LiDAR data. The linear regression model of a raster image is found by the contour recognition and it provides the classification accuracy as 83.91%. Non-rigid articulated object motion shows a periodic property, so this has been used as a strong cue for moving object classification. Optical flow is also very useful for object classification [2][26]. Residual flow can be used to analyze rigidity and periodicity of moving entities. It is expected that rigid objects would present little residual flow where as a non-rigid moving object such as a human being had higher average residual flow and even displayed a periodic component.

Texture based classification

In [7], the object is classified using the texture information extracted from the image. Every image captured from the camera is converted into Local Binary Pattern (LBP) as shown in fig 6. The LPB method is used to extract the corner feature points and detect the salient texture information with the capacity to tolerate the changes in the illumination. In [17], the cracks and squawks of the images are classified. The salient feature points are extracted from every LBP in order to identify the locations and corresponding sizes of the moving objects in the captured image frames. Once the feature points are extracted from the images a tracking algorithm can be applied to find the correspondences of those points between successive image frames. The tracking algorithm used in this paper is an optical flow approach. After determining the feature points from the LBP image it is obvious that the texture information of the image is extracted and the region based feature matching is performed by the obtained salient feature points extracted from previous LBP are compared to those features found from the current LBP with a block matching approach so that the corresponding feature points from the successive image frames can be identified. The images taken from the same scene under different illumination produces similar LBP images. The moving object segmentation is done to obtain the motion vectors for all of its salient feature points. Hence a clustering algorithm is used to segment all the motion vectors into different groups and each group represents a moving object. If multiple correspondences exist between feature points, the motion vectors of each feature points are then calculated to determine the best corresponding features on the current LBP. Finally, the clustering of motion vectors provides all the moving objects on image frames and the object can be successfully detected. The texture based classification provides 100% accuracy.

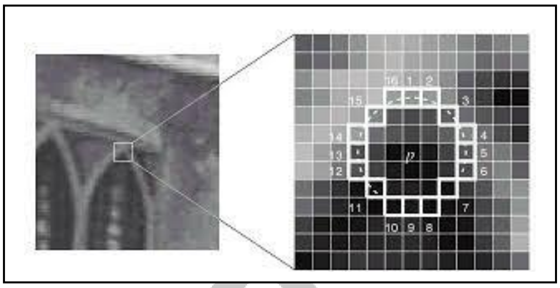

Fig. 6. Extracting LBP points

Color Based Classification

In [8], the images used for processing were taken from the camera located in the head of the robots, and the objects of interest to be detected were the actual AIBO robots. The detection of teammate and opponent robots is an interesting topic. In AIBO robot soccer, the robots have very limited field of view and the camera is placed on their movable head. The AIBO robots are white in color and similar to a dog and they use colored patches that do not cover the entire white body of the robot. This paper involves processing both the color segmented image and grayscale image taken by the robots. The use of color segmented images for producing initial hypotheses is for finding the location of robots in the image and grayscale images for final classification purposes. The initial hypotheses of the robots are obtained from the color segmentation algorithm by using the assumptions of field horizon, object not green and objects intersect field horizon. Now the object is classified using the grayscale image of the current vision frame, which is easily inferred from the YUV image returned by the robot hardware. Both the representations are combined together for the processing of a scene that allows each to make up for the deficiencies of the other, and provides high detection accuracy. The authors use a classifier to determine whether or not the detected objects are actually robots.

TABLEII. Comparison between different object classification algorithms

Algorithm	**Advantage**	**Disadvantage**
Feature Based Classification	➢ Useful in detecting the object when the shape is unavailable.	➢ Cannot be used for different environmental conditions.
Shape Based Classification	➢ Low computational time. ➢ Simple pattern matching.	➢ Unable to determine the object in dynamic situation. ➢ It provides moderate accuracy only.
Texture Based Classification	➢ High detection accuracy. ➢ Grayscale images are easy to process.	➢ High computational time.
Color Based Classification	➢ High detection accuracy.	➢ High computational time ➢ Colored images are difficult to process.

IV. *HAZE REMOVAL ALGORITHMS*

Dehazing Is The Process Of Removing The Fog Or Mist From The Image. The Following Is The Detailed Explanation Of The Dehazing Algorithms.

Color Attenuation Algorithm

In [22], the color attenuation prior algorithm is used to remove the haze in the image. This algorithm will initially create a linear model and the learning of parameters is done using supervised learning method. Here the hazy image is represented as follows.

$$I(x)=J(x)t(x)+A(1-t(x)), \text{ where } t(x)=e^{-\beta d(x)} \quad [22]$$

This equation is known as atmospheric scattering model and the terms I,J and A are the 3D vectors in the color space. The color attention prior algorithm provides high efficiency in haze removing process and it is free from over saturation problem. Using the color attention algorithm the original color of the image is preserved. The main drawback of this algorithm is the scattering coefficient β cannot be a constant in heterogeneous environment.

Dark Channel Prior Algorithm

In [27], the dark channel prior algorithm is used to remove the haze in the image. The dark channel algorithm will remove the haze by estimating the transmission and the transmission is refined using soft matting algorithm. The hazy image may contain some pixels with low intensity in any one of the color channel. These pixels are enhanced to recover the image. In [14], the DCP algorithm is only used with a guided filter. The guided filter estimates the improved transition map for better performance. The DCP algorithm is valid for the removal of heavy hazy image and provides good quality image even if the image is represented in gray scale with shadows. The drawback of this algorithm is it gets invalid when the scene object is similar to the airlight and also if there is no shadows.

TABLEIII. Comparison Of Haze Removal Algorithms

Algorithm	Advantage	Disadvantage
Color Attenuation Prior Algorithm	➢ High Efficiency. ➢ Free from over saturation. ➢ Maintains the original color of the image.	➢ Scattering coefficient β cannot be a constant in heterogeneous environment.
Dark Channel Prior Algorithm	➢ Valid for the removal of heavy hazy image. ➢ Valid for the removal of heavy hazy image. ➢ Works well even if the image is represented in gray scale with shadows.	➢ Invalid when the scene object is similar to the airlight and also if there is no shadows.

V. CONCLUSION

Thus the various object detection algorithms are analyzed which are used to detect an object using image processing. After detection of the object, the object is subjected to object classification algorithms. For this the various classification algorithms are also analyzed. In object detection, the diffusion based salient object detection method seems to have good advantages and in classification algorithm the texture based classification is found to be better than other classification algorithms. But both the algorithms possess high computational time. This can be reduced by using appropriate samples of training datasets without compromising the performance of the detection accuracy. The haze removal is done for providing better visibility to the object in the image.

REFERENCES

[1] Song Lu, Vijay Mahadevan, Nuno Vasconcelos, "Learning Optimal Seeds For Diffusion-Based Salient Object Detection", IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[2] Jianqin Yin Yanbin Han Wendi Hou Jinping Li Detection of the Mobile Object with Camouflage Color under Dynamic Background Based on Optical Flow" Advanced in control Engineering and Information Science Elsevier 2011.

[3] Zhao Liu, Hui Shen, Guiyu Feng, Dewen Hua, "Tracking objects using shape context matching", Neuro computing 83(2012),pp 47-55.

[4] Reza Oji, "An automatic algorithm for object recognition and detection based on ASIFT keypoints", Signal and image processing: An international journal, vol.3, no.5, October 2012.

[5] Juan Li, Chunfu Shao, Wangtu Xu, Hao Yue, "Real Time Tracking of Moving Pedestrians", 2009 International Conference on Measuring Technology and Mechatronics Automation, pp. 811-815.

[6] Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking Using Kalman Filter", International Journal of Computer Science and Mobile Computing,April 2013, pg.326 – 332.

[7] Yea-Shuan Huang, Zhi-Hong Ou, Hsiang-Wen Hsieh, Hung-Hsiu Yu," A Rapid Texture-Based Moving Object Detection Method",Proceedings of 2011 8th Asian Control Conference (ASCC).

[8] Juan Fasola, Manuela Veloso," Real-Time Object Detection using Segmented and Grayscale Images".

[9] Malathi. T, Manas Kamal Bhuyan, "Foreground Object Detection Under Camouflage Using Multiple Camera-based Codebooks", 2013 IEEE 2013 Annual IEEE India Conference (INDICON).

[10] Kinjal A Joshi, Darshak G. Thakore,"A Survey on Moving Object Detection and Tracking in Video Surveillance System", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.

[11] Huaizu Jiang, Jinjun Wang, Yihong Gong,"Online Multi-Target Tracking With Unified Handling of Complex Scenarios",IEEE Transactions On Image Processing, Vol. 24, No. 11, November 2015

[12] Michael Teutsch, Wolfgang Kruger," Spatio-Temporal Fusion of Object Segmentation Approaches for Moving Distant Targets".

[13] Weixing Wang, Lian Xu ," Retinex Algorithm on Changing Scales for Haze Removal with Depth Map ",International Journal of Hybrid Information Technology Vol.7, No.4 (2014),

[14] Zheqi Lin, Xuansheng Wang," Dehazing for Image and Video Using Guided Filter",Open Journal of Applied Sciences Supplement:2012 world Congress on Engineering and Technology

[15] Jianqin Yin, Yanbin Han, Wendi Hou, Jinping Li," Detection of the Mobile Object with Camouflage Color under Dynamic Background Based on Optical Flow", Advanced in control Engineering and Information Science Elsevier 2011.

[16] Munira A Jiwani, S.N.Dandare," Single Image Fog Removal Using Depth Estimation Based on Blur Estimation",International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013.

[17] Leontios J. Hadjileontiadis,"A Texture-Based Classification of Crackles and Squawks Using Lacunarity ",IEEE Transactions On Biomedical Engineering, Vol. 56, No. 3, March 2009.
[18] Kai-Fu Yang, Chao-Yi Li, Yong-Jie Li,"Multifeature-Based Surround Inhibition Improves Contour Detection in Natural Images ",IEEE Transactions On Image Processing, Vol. 23, No. 12, December 2014.
[19] Gemma Pons Bernad, Léonard Denise, Philippe Réfrégier, "Hierarchical Feature-Based Classification Approach for Fast and User-Interactive SAR Image Interpretation ",IEEE Geoscience And Remote Sensing Letters, Vol. 6, No. 1, January 2009.
[20] Xiang Zhang, Ce Zhu,"Camouflage Modeling For Moving Object Detection", 2015 IEEE
[21] Qingsong Zhu, Jiaming Mai, Ling Shao ,"A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior", IEEE Transactions On Image Processing, Vol. 24, No. 11, November 2015.
[22] Haifeng Hu,"Enhanced Gabor Feature Based Classification Using a Regularized Locally Tensor Discriminant Model for Multiview Gait Recognition",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 23, No. 7, July 2013.
[23] Peng Jiang, Nuno Vasconcelos, Jingliang Peng,"Generic Promotion of Diffusion-Based Salient Object Detection", IEEE Transactions.
[24] Belén Riveiro, Lucía Díaz-Vilariño, Borja Conde-Carnero, Mario Soilán, Pedro Arias," Automatic Segmentation and Shape-Based Classification of Retro-Reflective Traffic Signs from Mobile LiDAR Data",IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 9, No. 1, January 2016.
[25] SukHwan Lim, John G. Apostolopoulos, Abbas El Gamal, "Optical Flow Estimation Using Temporally Oversampled Video",IEEE Transactions On Image Processing, Vol. 14, No. 8, August 2005.
[26] Kaiming He, Jian Sun, Xiaoou Tang," Single Image Haze Removal Using Dark Channel Prior ".
[27] Bingquan Huo,Fengling Yin," Image Dehazing with Dark Channel Prior and Novel Estimation Model ",International Journal of Multimedia and Ubiquitous Engineering Vol. 10, No. 3 (2015), pp. 13-22.
[28] Shilpee A. Dave, Dr. M.S. Nagmode, Aditi Jahagirdar ,"Statistical Survey on Object Detection and tracking Methodologies", International Journal of Scientific & Engineering Research, Volume 4, Issue 3, March-2013 ISSN 2229-5518.
[29] Ariel Tankus, Yehezkel Yeshurun," Detection of Regions of Interest and Camouflage Breaking by Direct Convexity Estimation".