

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

## Framework for Characterization of Moving Objects

Julio-Alejandro Romero-González<sup>1</sup>, Ana M. Herrera-Navarro<sup>2</sup>, Hugo Jiménez-Hernández<sup>3</sup>

Ph.D. Student, Department of Computer Science, University Autonomous of Queretaro, Queretaro, Mexico<sup>1</sup>

Associate Professor, Department of Computer Science, University Autonomous of Queretaro, Queretaro, Mexico<sup>2</sup>

Associate Professor, Centro de Ingeniería y Desarrollo Industrial, Queretaro, Mexico<sup>3</sup>

**ABSTRACT:** Currently, interest in image recognition processes has gained a lot of importance, as technology has evolved and although there are robust vision applications, problems remain when it comes to analyzing and interpreting data perceived by machines dedicated to extracting information from images. These problems are due to the conditions of lighting, contrast, noise, environmental disturbances and change of position of objects, so the understanding of science has gained a lot of importance since many methods not only focus on segmentation are aimed at analyzing form, characteristics, geometry and physical issues that provide more information on the conditions of the object, based on this, this article presents a framework for the characterization of moving objects.

**KEYWORDS:** Text detection, Inpainting, Morphological operations, Connected component labelling.

### I. INTRODUCTION

Nowadays, interest in image recognition processes has gained a lot of importance, as technology has evolved and although there are robust vision applications, problems remain when it comes to analyzing and interpreting data perceived by machines dedicated to extracting information from images. These problems are due to the conditions of lighting, contrast, noise, environmental disturbances and change of position of objects, so the understanding of scenes has gained a lot of importance since many methods not only focus on segmentation. In which, segmentation is intended to subdivide an image to separate the parts of interest from the rest of the image, this subdivision depends on the problem to be solved, so there are different segmentation methods, of which some are based on characteristics such as colors, textures and borders to break down an image into uniform regions. Also, experts have carried out research for the purpose of reconnaissance of objects, they are aimed at analysis of form, characteristics, geometry and physical issues that provide more information about the conditions of the object, however there is still no general method for recognizing moving objects, for example, some authors detect people as isolated objects, but in real scenarios there are groups of people or couples, being the focus of this proposal a methodology is presented that allows to identify geometric structures of objects and classify them according to their geometric attributes.

### II. RELATED WORK

In [14] was designed an unsupervised algorithm to automatically adapt to soil conditions in different rooms where it assumes that the floor is the largest surface and has the same color and texture characteristics. In [11] was proposed a method based on learning, based on a multi-pooling architecture that takes the color of the image, encodes it according to its horizontal disparity, height above the ground, and angle the pixels and adds distance to the wall allowing the model to learn a more detailed representation of a scene. The method proposed in [10] is based on a number of binary trees where a pixel crosses the tree and the decoded path is based on a trait and a threshold stored in each node, once the pixel reaches a node, is assigned to the corresponding class distribution, the final result is obtained by averaging the distributions of the different nodes. Whereas the method in [2] is based on adapting the parameters of the segmentation

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

algorithm to environmental conditions, in combination with the border measurement for the assessment of segmentation, in order to reduce the calculation costs associated with the model during the adaptation phase. The segmentation through the fusion of color and depth of the objects presented by [9], models, the problem in a similar way to the standard segmentation routines as a random Markov field and is carried out by means of optimization of graphic charts, they determine pixel seeds in the first frame that are then automatically projected in the remaining frames and thus segment the objects in a scenario. Whereas, the algorithm proposed by [17] is based on the statistical approach for real-time background subtraction and shadow detection (SBGS) in which, the edge base and the Background intensity are modeled by SBGS. The color feature, edge feature and background intensity subtraction are integrated to achieve the best segmentation performance. Within this perspective, in [3] present a color and texture descriptor which is used in a two-stage rapid grouping framework using K-means to perform the segmentation of natural images online. The entire system works online along with location, 3D obstacle detection, and planning. The method proposed by [12] is based on multiple signals which are color, a direct line, edges, boundary points, PCs and HCM (Hue Co-occurrence Matrix). By combining these features, the image is segmented into several regions, as the mixture uses HCM features and methods, objects can be discovered when combining multiple predefined signals. On the other hand, the method used in [7] propose a new geocentric depth of images that encodes the height above the ground and the angle with gravity for each pixel, in addition to the horizontal disparity. These characteristics are used for learning convolutional neural networks.

For the method in [15], the main feature of this method is based on obtaining a package of past and future background images, which are compared and generate a median history for the pixels of the frames, detect changes between past and current frames to detect movement. Whereas, the propose in [18] was an algorithm by extending the ALM (Augmented Grange Multiplier) to exploit the information of the structure of the foregrounds and help to model the background, in addition to certain assumptions, such as low range or low-composition assumption (depending on whether pure background frames are provided) and formulate background extraction as a problem of matrix decomposition using regularization terms. On the other hand, in [19] was proposed a background subtraction system based on (Gaussian mixture modeling), in which each pixel dynamically searches for the best model, shared models are built for the background image and the foreground, small local movements (caused by noise) are discarded and a flexible mechanism for switching between the background and foreground models is added.

### III. BACKGROUND SUBTRACTION

Background subtraction applied to image sequences serves to detect the foreground; This means separate all the changes that occur in the foreground in order to detect movement, for this purpose different subtraction analysis techniques are used, of which only some of them will be explained.

#### *Image difference.*

This method consists in checking the difference between a reference image  $I_k$  and a new image  $I_{k+1}$  which will be the current image, the difference method is the common method of motion detection. This method adopts the pixel-based difference to find the moving object [16]. This method is very sensitive to the  $Th$  threshold.

$$Th < |I_k - I_{k+1}| \quad (1)$$

#### *Temporal median filter.*

This method consists in estimating the background by accumulating  $n$  images and calculating the average or median for each pixel [5]. This method is quite fast, but it consumes a lot of memory which is determined based on the following equation:

$$memory = n * size(image) \quad (2)$$

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

To determine the average of background:

$$B_{(i+1)} = \alpha * F_i + (1 - \alpha) * B_i \quad (3)$$

Where B is the background image, F the foreground image,  $\alpha$  is constant of learning (0.05) and i denotes the previous iteration.

### Mixture of Gaussians.

It is an algorithm that represents the probability distribution of a pixel if observed and part of an estimate  $\theta$  until the convergence is detected, in each iteration, each pixel is modeled with a Gaussian, from which the maximum probability is determined to calculate the parameters of the model:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (4)$$

Where P is the probability of observing,  $\omega$  the weight associated with the Gaussian i. Each pixel of the model is continuously updated and trained, as it is updated, Gaussian components must be standardized [4].

$$\omega_{i,t} = (1 - \alpha)\omega_{i,t} + \alpha \quad (5)$$

$$\mu_{i,t} = (1 - \rho)\mu_{i,t} + \rho X_t \quad (6)$$

Where  $\alpha$ ,  $\rho$  are learning constants and  $\Sigma$  is the covariance matrix.

## IV. FEATURE DETECTORS

Obtaining feature points of an image is intended to provide a representation of an object, so it is sought that these features are invariant to rotation, scale or lighting effects [1], divide the main methods for feature extraction in four main groups, based on general detectors of: edges, corners, binary objects and regions. For the purposes of this article, the Harris, Laplacian of Gaussian and Harris-Laplace corner detectors are explored.

### Harris detector

The Harris operator was developed by [8], it is used in the literature to infer characteristics of an image, its rooted in the invariance to scale, rotation and lighting changes. Feature points found by Harris are used to decrease the amount of information that is processed, Harris's algorithm is defined as follows:

- i. Calculate the gradient of image as equation 6.

$$G = \sqrt{I_x^2 + I_y^2} \quad (6)$$

Where  $I_x$  and  $I_y$  represents change of intensity on axis x and y respectively.

- ii. Calculate products of derivative of each pixel.

$$G_{x^2} = G_x \cdot G_x \quad (7)$$

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

$$G_{y^2} = G_y \cdot G_y \quad (8)$$

$$G_{xy} = G_x \cdot G_y \quad (9)$$

iii. Calculate sums of derivative products at each pixel.

$$S_x = \sum G_{x^2} \quad (10)$$

$$S_y = \sum G_{y^2} \quad (11)$$

$$S_{xy} = \sum G_{x^2} \quad (12)$$

iv. Now it is defined the Hessian matrix

$$H(x, y) = \begin{bmatrix} S_x(x, y) & S_{xy}(x, y) \\ S_{xy}(x, y) & S_y(x, y) \end{bmatrix} \quad (13)$$

v. Finally, the detector response is calculated for each pixel

$$R = \|H\| - k \cdot (\text{tr } H)^2 \quad (14)$$

### Laplacian of Gaussians

The Laplacian of Gaussians (LOG) filter is an operator that finds the change of intensity in all directions, in this method the application of a Gaussian filter is combined to decrease the sensitivity to noise at the time and a Laplacian filter, this helps to detect the points of interest in an image highlighting the edges [13]. The LOG algorithm works in space scale his representation is defined as follows:

$$L(x, y, \sigma) = g(x, y, \sigma) * f(x, y) \quad (15)$$

Where  $g(x, y, \sigma) = \frac{1}{2\pi\sigma} \cdot e^{\left(-\frac{x^2+y^2}{2\sigma}\right)}$ , the parameter  $\sigma$  represents de standard deviation, which is the distribution of gaussian function over x and y axis.

For resolve the problem with the object scale a multiscale detection its applicated, for this a gradient its calculated.

$$\nabla^2 L(x, y, k\sigma) = L_{xx}(x, y, k\sigma) + L_{yy}(x, y, k\sigma) \quad (16)$$

Finally, keypoints are calculated by obtaining the difference between values positive maximum and the negative minimum.

$$\text{Keypoints} = \text{argmaxmin}(\nabla^2 L(x, y, k\sigma)) \quad (17)$$

### Harris Laplace Detector

This algorithm searches for corners in multi-scale images and uses the place operator to detect the appropriate scale that helps measure the similarity between the Harris operator and the image.

$$L = \det \left( \begin{bmatrix} L_x^2(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_y^2(x, \sigma) \end{bmatrix} \right) - k * \text{Trace}^2 \left( \begin{bmatrix} L_x^2(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_y^2(x, \sigma) \end{bmatrix} \right) \quad (18)$$

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

## V. EXPERIMENTAL RESULTS

In this section the experimental results of the framework are presented, the principal experiments consist of adjusting the parameters of the background model of temporal means and characterizing the light changes associated with the texture of the objects that are in motion. In order to demonstrate the effectiveness of this proposal, three sequences of 700 images have been analyzed, dimensions of images of the first sequence are 266 x 342 pixels, dimensions of the second sequence of 479 x 354 pixels and dimensions of images of the third sequence are 565 x 581 pixels. Sequences have been obtained from the database used in [6]. The results of those experiments are showed in fig. 1, fig. 2 and fig. 3.



Fig. 1. Features detected in the first scenario by a) Harris detector, b) Harris-Laplacian detector and c) Laplacian of Gaussians detector. Source: Own elaboration.

In fig 1, it can be observed features from Harris, Harris-Laplacian and LOG algorithms. In which Harris-Laplacian show more features from object to difference from Harris and LOG which are focused on edges. Parameters used are  $\alpha = 0.3$ ,  $\sigma = 2$  for background model;  $\sigma = 2$ ,  $k = 1.2$  for the Harris-Laplacian algorithm,  $k = 0.04$  for Harris algorithm and finally  $\sigma = 0.7$  and  $k$  values from 1 to 5 for LOG algorithm.

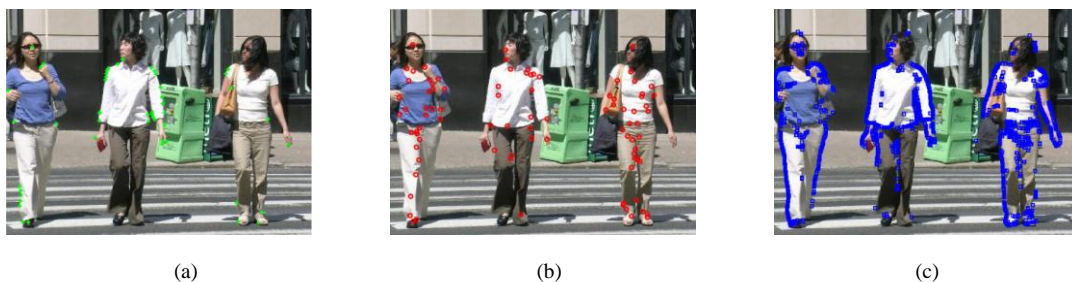


Fig. 2. Features detected in the second scenario by a) Harris detector, b) Harris-Laplacian detector and c) Laplacian of Gaussians detector. Source: Own elaboration.

In fig 2, LOG shows more features in relation to objects in movement, these features can be excessive because it has described the intensity changes from edges mainly. Harris detected less features, but they are more representative of the objects than LOG. In spite of changes of illumination and the cloth brightness Harris-Laplacian algorithm has the best features of objects. Parameters used are  $\alpha = 0.7$ ,  $\sigma = 2.5$  for background model;  $\sigma = 1.5$ ,  $k = 1.2$  for the Harris-Laplacian algorithm,  $k = 0.04$  for Harris algorithm and finally  $\sigma = 0.5$  and  $k$  values from 1 to 5 for LOG algorithm.



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

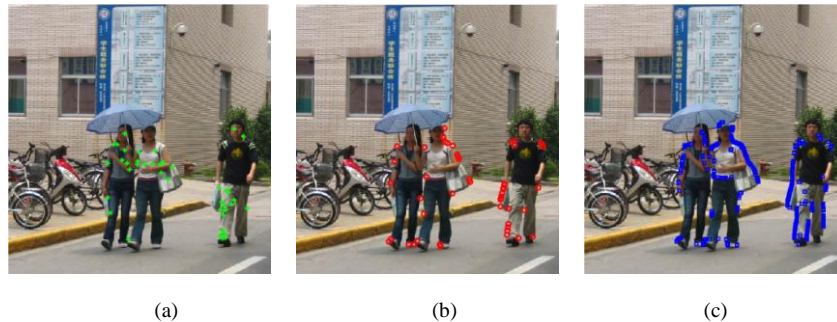


Fig.3.Features detected in the third scenario by a) Harris detector, b) Harris-Laplacian detector and c) Laplacian of Gaussians detector. Source: Own elaboration.

The third scenario shows some stationary objects that can be observed as a movement. In fig 3, it can be observed features detected from algorithms. Parameters used in this scenario are  $\alpha = 0.3$ ,  $\sigma = 1.5$  for background model;  $\sigma = 2$ ,  $k = 1.2$  for the Harris-Laplacian algorithm,  $k = 0.04$  for Harris algorithm and finally  $\sigma = 0.5$  and  $k$  values from 1 to 5 for LOG algorithm. In order to quantify the information obtained by algorithms, statistical measures were used to measure the performance and precision from algorithms, whereby ROC curves were calculated using sequence of 300 images from three new scenarios, results are shown in fig 4.

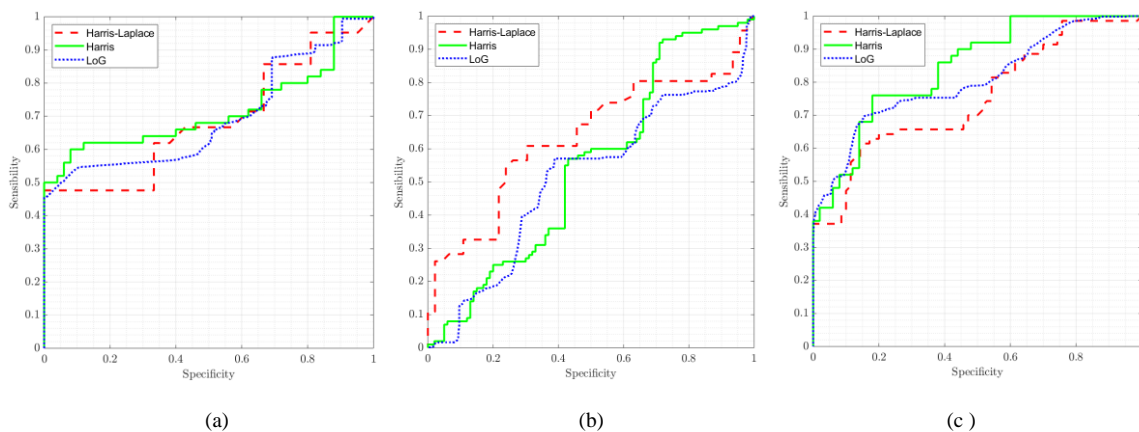


Fig.4.ROC curves. This figure shows the performance of Harris detector, Harris-Laplacian detector and Laplacian of Gaussians detector. Source: Own elaboration.

In Fig. 4, sensitivity-precision curves are shown, in which the Harris-Laplace algorithm has shown better performance and more relevant information is considered in spite of lighting conditions, contrast, noise, environmental disturbances and change of position of objects, reflections and shadows. In scenarios (a) and (c), algorithms evaluated have good results, but in (b) only Harris-Laplace has better positive predictions than others.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 9, Issue 1, January 2020

## VI. CONCLUSION

This article presents the development of a framework for the description of scenes, which by means of background subtraction models, moving objects of static objects are segmented, so to extract features of the region of interest, descriptors are used generally to analyze its geometry, understand its dynamics and the conditions of the object. The framework presented, has a wide use since algorithms can be developed for detection of hostile situations, automatic detection of people or auto detection, trajectory tracking, despite the conditions of changes in the position of the object, changes in lighting or noise.

## REFERENCES

- [1] Bernal, J., Vilarino, F., and Sanchez, J., "Feature Detectors and Feature Descriptors: Where We Are Now", Review of International Studies, vol 1, pp. 112, 2010.
- [2] Bhanu, B., and Peng, J., "Adaptive integrated image segmentation and object recognition", IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, vol 30(4), pp. 427–441, 2000. <http://doi.org/10.1109/5326.897070>
- [3] Blas, M. R., Agrawal, M., Sundaresan, A., and Konolige, K., "Fast color/texture segmentation for outdoor robots", 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, vol. 1(d), pp. 4078–4085., 2008. <http://doi.org/10.1109/IROS.2008.4651086>
- [4] Bouwmans, T., El Baf, F., and Vachon, B., "Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey", Recent Patents on Computer Science, vol. 1(3), pp. 219–237, 2008. <http://doi.org/10.2174/2213275910801030219>
- [5] Chowdhury, A., Cho, S. J., and Chong, U. P., "A background subtraction method using color information in the frame averaging process", Proceedings of the 6th International Forum on Strategic Technology, IFOST 2011, vol. 2, pp. 1275–1279, 2011. <http://doi.org/10.1109/IFOST.2011.6021252>
- [6] Dollar, P., Wojek C., Schiele, B., and Perona, P., "Pedestrian detection: A benchmark", 2009 IEEE Conference on Computer Vision and Pattern Recognition, vol 1, pp. 304–311, 2009.
- [7] Gupta, S., Girshick, R., Arbeláez, P., and Malik, J., "Learning rich features from RGB-D images for object detection and segmentation", Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7, pp. 345–360, 2014. [http://doi.org/10.1007/978-3-319-10584-0\\_23](http://doi.org/10.1007/978-3-319-10584-0_23)
- [8] Harris, C., and Stephens, M., "A combined edge and corner detector", in Proceeding of the 4th Alvey Vision Conference, pp. 189–192, 1988.
- [9] He, H., and Upcroft, B., "Automatic object segmentation of unstructured scenes using colour and depth maps", IET Computer Vision, vol. 8(1), pp. 45–53, 2014. <http://doi.org/10.1049/iet-cvi.2013.0018>
- [10] Hermans, A., Floros, G., and Leibe, B., "Dense 3D semantic mapping of indoor scenes from RGB-D images", Proceedings - IEEE International Conference on Robotics and Automation, vol. 1, pp. 2631–2638, 2014. <http://doi.org/10.1109/ICRA.2014.6907236>
- [11] Husain, F., Schulz, H., Dellen, B., Torras, C., and Behnke, S., "Combining Semantic and Geometric Features for Object Class Segmentation of Indoor Scenes", IEEE Robotics and Automation Letters, vol. 2(1), pp. 49–55, 2016. <http://doi.org/10.1109/LRA.2016.2532927>
- [12] Kim, D.-N., Trinh, H.-H., and Jo, K.-H., "Regions segmentation using multiple cues for robot navigation on outdoor environment", 2007 International Conference on Control, Automation and Systems, vol. 1, pp. 1768–1772, 2007. <http://doi.org/10.1109/ICCAS.2007.4406631>
- [13] Kong, H., Akakin, H. C., and Sarma, S. E., "A generalized laplacian of gaussian filter for blob detection and its applications", IEEE Transactions on Cybernetics, vol. 43(6), pp. 1719–1733, 2013. <https://doi.org/10.1109/TSMCB.2012.2228639>
- [14] Ma, L., Wang, J., Bo, Z., and Wang, S., "Automatic floor segmentation for indoor robot navigation", ICSPS 2010 - Proceedings of the 2010 2nd International Conference on Signal Processing Systems, vol. 1, pp. 684–689, 2010. <http://doi.org/10.1109/ICSPS.2010.5555399>
- [15] Rahaman, A., Hasan, M. M., Ahmed, R., Maswood, M. S., and Rahman, M. M., "A new background updating model for motion detection considering future frame", 16th Int'l Conf. Computer and Information Technology, vol 1, pp. 378–382, 2014. <http://doi.org/10.1109/ICCITech.2014.6997296>
- [16] Singla, N., "Motion Detection Based on Frame Difference Method", International Journal of Information and Computation Technology, vol. 4(15), pp. 1559–1565, 2014.
- [17] Siricharoen, P., Aramvith, S., Chalidabhongse, T. H., and Siddhichai, S., "Robust outdoor human segmentation based on color-based statistical approach and edge combination", 1st International Conference on Green Circuits and Systems, vol. 1, pp. 463–468, 2010. <http://doi.org/10.1109/ICGCS.2010.5543017>
- [18] Xin, B., Tian, Y., Wang, Y., and Gao, W., "Background Subtraction via generalized fused lasso foreground modeling", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 4676–4684, 2015. <http://doi.org/10.1109/CVPR.2015.7299099>
- [19] Chen, Y., Wang, J., and Lu, H., "Learning sharable models for robust background subtraction", 2015 IEEE International Conference on Multimedia and Expo (ICME), vol. 1, pp. 1–6, 2015.