A Proposed Security System Using Object Proposals via Superpixel Segmentation

Vikas Upadhyaya

Assistant Professor, Electronics & Communication Engineering area
NIIT University, Neemrana
vikas.upadhyaya@niituniversity.in

Abstract**:** ***A simple security system is presented in this paper, which can be implemented through multiple surveillance cameras, to detect trespasses and breaches of security in a wide range of real-world applications, especially on properties without fencing. The main distinguishable factor contained in the paper is the developed proposed algorithm, which largely deals with grouping objects based on superpixel properties and subsequently proposing objects in a given image. Some of the basic concepts explored include superpixel segmentation (with the help of the concept of SEEDS- Superpixel Extraction using Energy Driven Sampling) and a logical grouping of superpixels proposed to be in the same object.***

Keywords**: Image segmentation, object proposals, Object Proposals Using Superpixel Segmentation (OPUSS), Superpixel Extracted by Energy Driven Sampling (SEEDS), Security System.**

(Article history: Received: March 2022 and accepted May 2022)

I. Introduction

Criminal acts of behavior pose a serious threat to individual as well as communal well-being. Anti-social activities disrupt the harmony of everyday life. The recent spike in violent crimes calls for the need of more robust and efficient security systems. This security system has been developed keeping in mind simplicity of approach in both hardware and software. In order to identify and track unwanted objects, superpixel grouping has been used. Object proposals have been done though an algorithm with a new approach- OPUSS. The hardware module is used to relay breach of security to the owner/concerned authorities.

The algorithm developed (OPUSS) to make object proposals has the following characteristics:

1. Detecting the presence of multiple objects in the foreground.
2. Be able to differentiate between two objects in close proximity.
3. Be able to track the object in order to detect motion in successive frames.

The requisites mentioned above, are however pivotal upon the fact that these tasks be done so as to reduce the execution speed and memory storage requirements.

In the domain of digital image processing, image segmentation is a transition from low level image processing to processing with more complex attributes, features and scenes, and often the success of image analysis depends on a reliable and robust segmentation process. For the segments to be useful, they must be strongly related to regions/objects of interest in the image but in reality, accurate and reliable segmentation is a challenging problem. Therefore, the ability of OPUSS to make successful object proposals and its universality were the most important parameters to consider. In order to support this claim, a qualitative analysis obtained through different results has been presented in the research paper.

A key concept used in this system is superpixel segmentation. Superpixels can simply be defined as a collection of pixels that are similar in color and texture. The main advantage of superpixels in this scenario is that it completely eliminates the need of analyzing individual pixels and thus simplifies further operations on the output of the segmentation process.

The proposed hardware which is used for the alert system consists of a GSM module which is connected to the controlling PC using a USB to TTL converter.

The subsequent sections in the research paper discuss the various sub-processes of the proposed system. Details of the object detection and tracking process have been elaborated in Section 3. Section 4 deals with the proposed alert system and its hardware implementation followed by results and conclusion in Section 5 and 6, respectively.

II. EXISTING SOLUTIONS

Previous surveys and the number of conferences and papers on image segmentation, object proposal and detection show that it is a highly researched area.

In the simple background subtraction approach described in [2], the distribution of the pixels related to any object was estimated using a density function. In order to estimate the belongingness of the pixel to foreground or background, a probabilistic approach was used. This method was relatively easy in implementation but was found to be computationally expensive.

In the approach described in [3], Gaussian distribution was used to model any pixel in the given image. This method was different to other tradional methods where only one type of distribution was preferred to model the pixel distribution in the given image. Depending on the persistence and variance values in the Gaussian function, the approach facilitates the classification of the pixels to foreground or background region.

The approach described in [4] takes a different direction by using multiple classifiers to track any object. Defined classifiers covers different scales of orientation. The classifiers attempt to identify the presence of the object in the given image at any specified size. The method helps in identifying the object in the given image where the object may be of different orientation.

The main consideration while working on this research work was to develop a method suited for colored images. Graph based methods discussed in [4] under Superpixels is a versatile area which is often used for Segmentation in Digital Image Processing. For the detection of generic objects, another technique is proposed in [5] where authors suggested to use partial spanning trees of superpixels. This helps in resolving the conflict of sampling connected groups of superpixels. These superpixeles may belong to the same object. The process is implemented using Randomized Prism’s algorithm.

The optical flow method as described in [6] defines optical flow as the movement of different objects with different velocities and represent them as the movement of brightness. The authors also suggest that the optical flow may exist due the relative motion between the objects and the stationary observer. The method is superior in the detection of slow movement of objects in any frame.

Other segmentation approach as described in [7] propose the application of greedy algorithm for the detection of similarities between the region of interest and the near by region in any image. The algorithm combines the similar regions and later find the similarity as suggested above. The proposed method is used in object recognition using selective search. In order to get the optimum results, the segmentation process is performed in multiple color channels.

In the domain of human detection, Papageorgiou et al [8] describe a pedestrian detector based on a polynomial SVM using rectified Haar wavelets as input descriptors, with a parts (sub-window) based variant in [9].

Another important consideration was successful segmentation of the given image. However, segmentation-heavy approaches are generally computationally expensive. State-of-the-art segmentation methods require anything between half a minute up to several minutes to compute segmentation proposals for a single image. RIGOR, a computationally efficient segmentation method described in [11], where a precomputation step is employed which reuses computation for multiple related parametric min-cuts and also avoids lengthy computations by using a set of choices from a segmentation pool. The RIGOR approach improves upon the graph cutting technique.

An approach for image segmentation using weighted graphs in black and white images is described in [12] using the two cues of color and texture. Texture is measured using textons and this approach was modelled on the normalized graph cut approach.

III. OBJECT DETECTION AND TRACKING

The system proposes, detects and tracks objects by means of the following processes:

A. Superpixel Extraction using Energy Driven Sampling (SEEDS)

Image segmentation partitions a given image into multiple regions (or, in this case into groups of pixels) based on a predefined attribute or constraint. The superpixel segmentation process used in OPUSS is based on the method used in SEEDS.

The concept of SEEDS is based on updating the segmentation grid in the form of blocks (after every iteration), as opposed to other methods that progressively build the superpixel region. Instead of incrementally building the superpixels by adding cuts or growing superpixels, SEEDS iteratively refines a superpixel partitioning by exchanging pixels between neighboring superpixels.

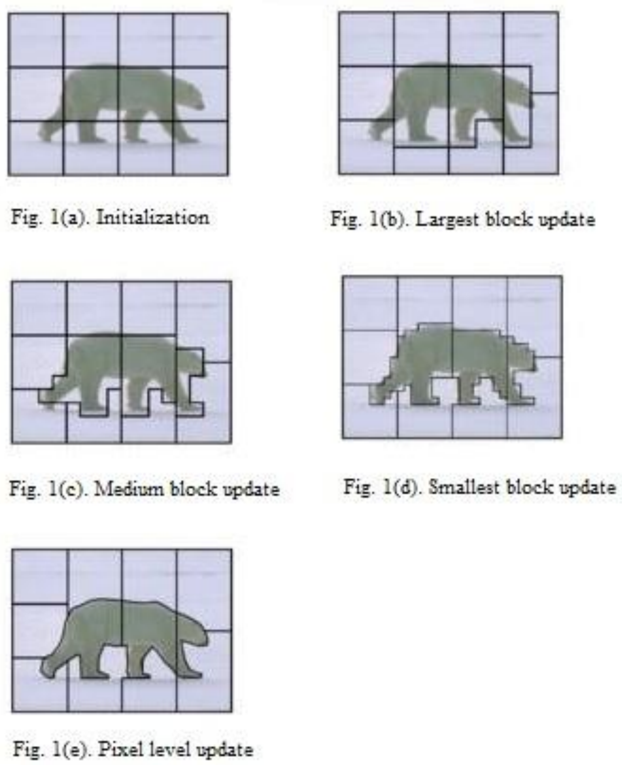

Fig. 1. Superpixel Segmentation

In Fig 1(a) the initialization of the image is shown. The superpixel segments are then updated on a largest to smallest block basis as seen in (b), (c) and (d). Finally, the stray pixels that do not belong to any segment are integrated into the nearest neighbor.

The pixel exchange is based on the homogeneity of the color distribution of the superpixels and a term that encourages smooth boundary shapes.

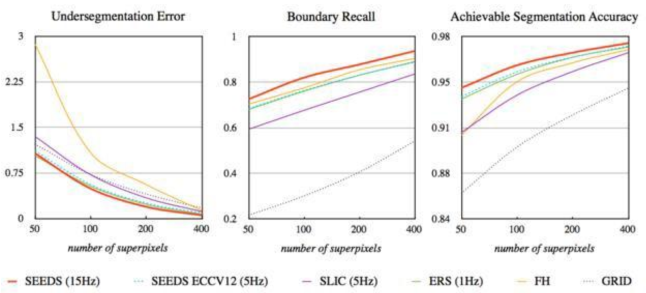

Fig. 2. Comparison between different techniques

It can be observed that SEEDS (represented in red) performs better than most existing technologies such as SLIC, ERS, FH etc.

Detection of the object becomes difficult if there exist high degree of similarity between the color of the object and background. Similarly, lighting and illumination conditions also affect the performance of the proposed object detection process. To handle lighting and illumination issues, the camera must be chosen to capture better images in the poor lighting conditions.

B. Region Adjacency Graph

A Region Adjacency Graph (RAG) is constructed by defining a vertex per object and joining two vertices with an edge if the objects are adjacent. RAGs can be computed for labelled 2D or 3D image. Two regions are considered as neighbor if they are separated by a small number of pixels in the horizontal or vertical direction. The output of the segmentation process is stored in a 2D array which contains the labels given to each segment. Two regions are considered as neighbors if they are separated by a small number of pixels in the horizontal or vertical direction, as well as depth direction for 3D images

In order to determine connectivity, the RAG is implemented using the average RGB values of each segment (as calculated in the segmentation process). This value is updated as the center of the superpixel. The difference between two such center values is then updated as the edge between two superpixels.

The main motive behind this process is the fact that any two superpixels belonging to the same object will have similar color and texture values. Thus, two similar superpixels are connected by a low weight edge and two dissimilar superpixels are connected by a high weight edge. This is translated into logic as; the value of the edge will be less than a predefined threshold only if two connected superpixels belong to the same object. Hence, we can eliminate the high edge weights to be left with only those superpixels pertaining to the objects.

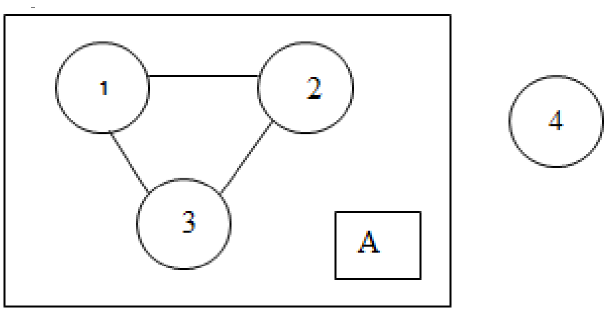

Fig. 3. Weight assignment of the connect superpixels using RAG concept

Using the RAG concept as indicated in Fig. 3, the weight of the edge connective superpixels 1 to 2, 2 to 3, and 3 to 1 is 3, 2 and 5 respectively. The weight of the edge connecting superpixel 4 to 1, 2 and 3 is 27, 22 and 24 respectively. If the first threshold is set at 10 (i.e. the maximum edge weight between two superpixels belonging to one object), then superpixel 4 will not satisfy the criteria and will not be therefore considered to be a part of the object, A.

C. Object Proposals

TABLE 1. DEFINED PARAMETERS FOR MAKING OBJECT PROPOSALS

Parameter	Definition
n_i	Number of superpixels being considered for the first threshold
n_f	Number of superpixels satisfying the first threshold
T_1	Primary threshold defined for cutting RAG edges based on edge weight
T_2	Maximum number of superpixels that can belong to one object

Table 1. lists all the parameters used for making object proposals through logical elimination

The algorithm employs a graph cutting technique to eliminate those edges of the RAG which do not satisfy the first threshold. Defining e_i as the value of the weight for the ith edge, if n_i superpixels are taken, then only those superpixels will be a part of the final object proposal which satisfy:

$$e_i =< T_1, i= 1, 2, 3\ldots..k \qquad (1)$$

In any given image, this condition will yield a wide range of object superpixels. The question then arises as to how one can differentiate between useful and non-useful proposals.

A simple way to ensure that the output contains only useful object proposals is to define another hard threshold T_2. Therefore, given that condition (1) is true; another check is performed, which can be represented as:

$$n_f < T_2 \qquad (2)$$

where n_f is the number of superpixels satisfying condition T_1.

This prevents the proposals of large background components as objects of interest (such as the sky, trees, buildings etc.)

At this stage, OPUSS may produce some false positives (constituents of the image that are not objects but are proposed as such). These false positives can be discarded in machine learning systems to which we can directly feed the output of OPUSS.

D. Object Tracking

Object tracking is achieved by running the algorithm on individual frames. Once the objects are detected in any frame, a bounding box is placed on them for the tracking purpose. The proposed bounding box can be placed using the inbuilt function available in the library. The camera frame rate used in this research was 25 frames per second but the same can be tweaked according to the hardware processing capability of the system.

IV. PROPOSED ALERT SYSTEM

The hardware of the alert system consists of the SIM900 GSM module by SIMCom. The module features an industry-standard interface, and delivers GSM/GPRS 850/900/1800/1900MHz performance for voice, SMS, Data, and Fax in a small form factor and with low power consumption. To connect the module to a PC, a USB to TTL converter is used. The chosen components are commonly available, easy to use and adhere to industry standards.

A mask is used which defines the restricted area in the video feed. This mask is logically ANDed with the object predictions, therefore if an object is present in the defined restricted area, an alert is generated which triggers the GSM module to send an SMS to a predefined mobile number. The GSM module is used to send an SMS using the AT Cellular Command set.

V. RESULTS

A. OPUSS Optput:

Superpixel segmentation (SEEDS), Region Adjacency Graph and Object Proposal:

Fig. 4(a). Input to OPUSS

Fig. 4(b). Output of SEEDS

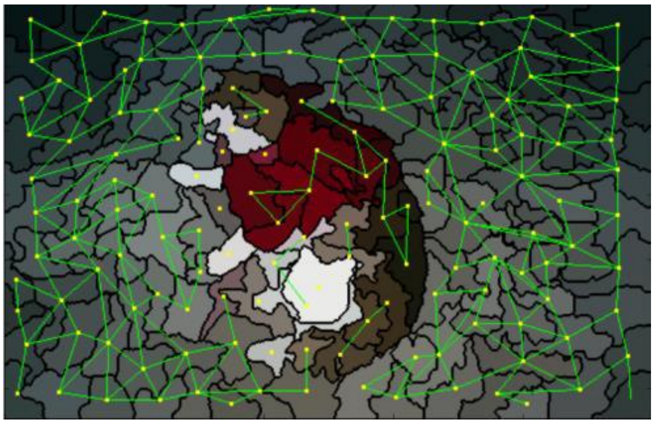

Fig. 4(c). Construction of RAG

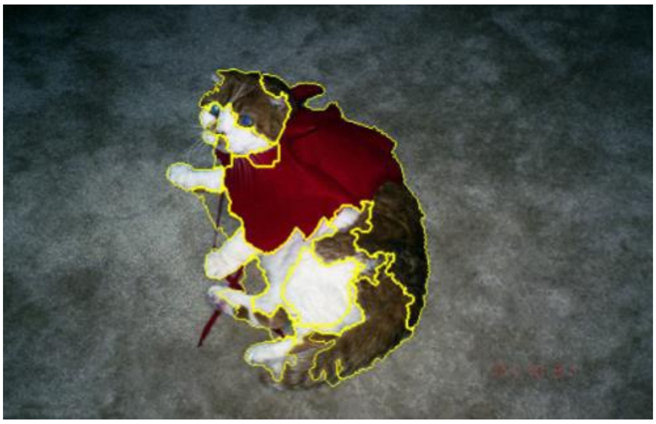

Fig. 4(d). Final Object proposal

At the end of the segmentation stage, each superpixel is assigned a certain color, which is the RGB average of the RGB value of each pixel in the superpixel. We obtain approximately 200 superpixels from SEEDS.

Each superpixel is considered to be a vertex in the graph. The RAG is constructed on the output of SEEDS and at this stage two thresholding processes occur. First, superpixels with connecting edge weights less than the predefined threshold are added to the same group; secondly, a group of superpixels is only proposed as an object if the number of superpixels in the group is less than a predefined threshold.

B. *Object Tracking*

Once the object proposals have been obtained, they are tracked by means of bounding boxes as visualized in Fig. 5.

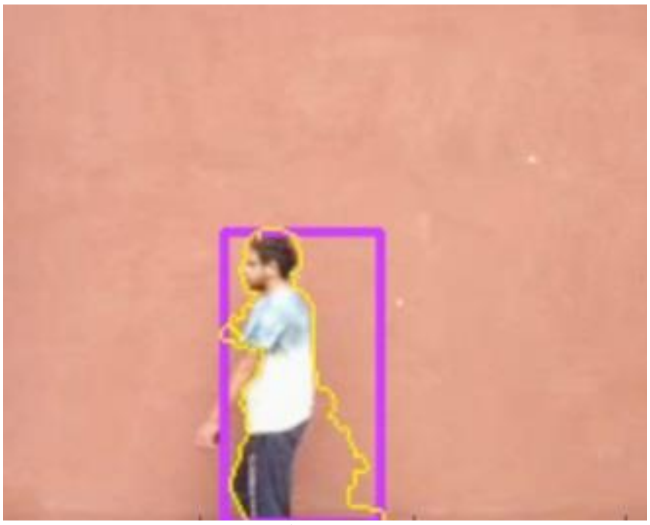

Fig. 5. Placement of Bounding box on the object

As shown in Fig. 5, Bounding boxes have been placed using the corner points of the superpixels in the proposed objects. To remove duplicates, a secondary check is done so that each object is bound by a single box only and multiple boxes are eliminated.

C. *Integration Into the Proposed Alert System*

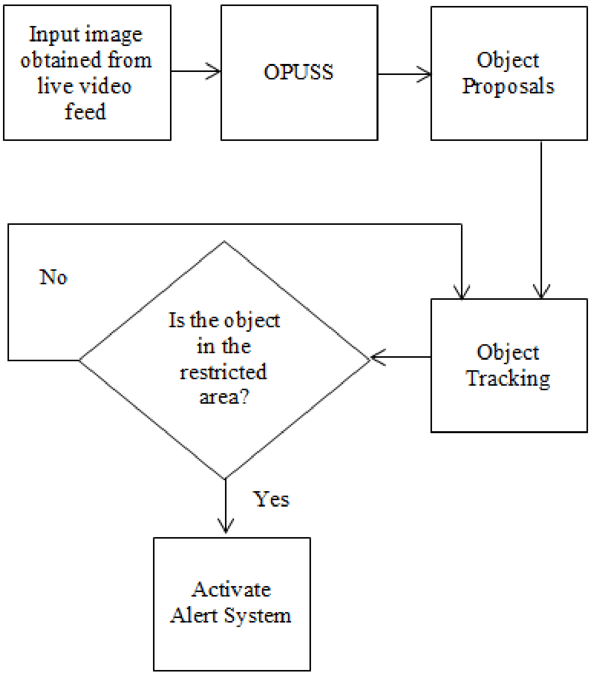

Fig. 6. A flowchart representing the working mechanism of the system.

This application presented in the subsequent paragraphs has been developed specifically for outdoor spaces without fencing. In order to detect intrusions in such properties, a restricted region is first defined, around the property.

Implementation:

Fig. 7(a). Input to the system

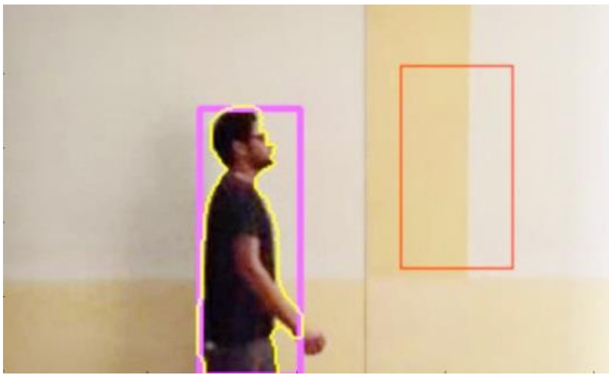

Fig. 7(b). Detected object

The process described in the flowchart depicted in Fig. 6, has been visually realized in Fig. 7. The input to the object proposal process, depicted in Fig. 7(a) is segmented and the detected object is presented in Fig. 7(b). The restricted region is defined by the bounding box in red in Fig. 7(b).

Once the object enters the restricted region, the hardware system sends an alert (by means of a text message) to the owner/concerned party.

VI. CONCLUSION

The security system proposed has been developed based on image processing concepts. This approach provides a strong advantage while making improvements to the system since the image segmentation method can be improved and made more robust through future work and advancements. By creating ground truth references for all possible objects, machine learning tools can be used to train the algorithm to better detect intrusions.

Although the current algorithm is able to successfully propose objects in an image, there is still room for further optimization in order to reduce the presence of false positives. While the results can indeed be further refined through extensive quantitative analysis, the current model produces satisfactory results when tested.

ACKNOWLEDGMENT

Jai Verma and Sinjini Mitra who were the students of NIIT University contributed to this research work during their BTech program.

REFERENCES

[1] V. den Bergh M., Boix X., Roig G., de Capitani B. and Van Gool L., “SEEDS: Superpixels Extracted via Energy-Driven Sampling”,in *European Conference on Computer Vision*, vol. 7, pp. 13-26, 2012.

[2] Ahmed Elgammal et al., “Non-parametric Model for Background Subtraction,” in *6th European Conference on Computer Vision*, Dublin, 2000.

[3] Chris Stauffer, W. Eric L. Grimson, “Learning patterns of activity using real time tracking”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 22, Issue: 8, Aug 200.

[4] Henry Schneiderman and Takeo Kanade, “Object Detection Using the Statistics of Parts”, International Journal of Computer Vision, Vol. 56; Iss. 3, 2004.

[5] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk, “SLIC Superpixels”, *EPFL Technical Report 149300*, June 2010.

[6] Santiago Manen, Matthieu Guillaumin, Luc Van Gool, “Prime Object Proposals with Randomized Prim’s Algorithm”, in *International Conference on Computer Vision*, Sydney, NSW, Australia, 2013.

[7] B.K.P. HORN AND B.G. SCHUNCK, “Determining Optical Flow”, *Artificial Intelligence*, vol. ED-17, pp. 185-203, 1981.

[8] A. Mohan, C. Papageorgiou, and T. Poggio. “Example-Based Object Detection in Images by Components,” *PAMI*, vol. 23, no. 4, pp. 349–361, April 2001.

[9] C. Papageorgiou and T. Poggio, “A Trainable System for Object Detection”, *International Journal of Computer Vision*, vol. 38, pp. 15–33, 2000.

[10] Koen E. A. van de Sande, Jasper R. R. Uijlings, Theo Gevers, Arnold W. M. Smeulders, “Segmentation as Selective Search for Object Recognition”, in 2011 *IEEE International Conference on Computer Vision*, Barcelona, Spain, 2011.

[11] Ahmad Humayun et al., “RIGOR: Reusing Inference in Graph Cuts for generating Object Regions”, in 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014.

[12] Jitendra Malik, Serge Belongie, Thomas Leung and Jianbo Shi, “Contour and Texture Analysis for Image Segmentation”, International Journal of Computer Vision, vol. 43, no. 1, pp. 7–27, 2001.

AUTHOR PROFILE

Dr. Vikas Upadhyaya (PhD)
Assistant Professor, NIIT University.
Having experience of more than 15 years in academics and specialized in the area of VLSI, Computer Vision and Sensors.
Has published more than 15 research papers in reputed Journals and Conferences.