Pak. J. Biotechnol. Vol. 14 (Special Issue II) Pp. 372-376 (2017) Gautam, K.S. and S.K. Thangavel

# HIDDEN OBJECT DETECTION FOR CLASSIFICATION OF THREAT

[1]K.S. Gautam, [2]Senthil Kumar Thangavel

Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham Amrita University, India

cb.en.d.cse14006@cb.students.amrita.edu[1], t_senthilkumar@cb.amrita.edu[2]

*ABSTRACT*

The paper proposes an intelligent K-means segmentation algorithm that clearly segments foreground objects and completely occluded objects. When a person completely occludes an object while entering into the area of video surveillance, it is considered as an anomaly. The paper comes up with a robust technical solution to address this. The proposed algorithm chooses an optimal value for K and segments the object. The scope of the system extends to the area such as prison, airport etc.. where there is a need to monitor completely occluded objects and other objects in the foreground. The system is tested with images from Stereo Thermal Dataset and achieves a precision rate of 88.89% while segmenting objects. From the experimental results, we infer that the proposed algorithm is robust in segmenting the objects without losing its shape and number.

Keywords— Video Analytics, Occlusion, Segmentation, Motion vector optimization, Object Detection, Thresholding.

## I. INTRODUCTION

Addressing a problem with a technical solution is comparatively reliable than handling it with manual supervision. It is an open secret that the advent of debit cards and credit cards has minimized robbery and pickpockets. Likewise in video surveillance, though the actions of people are manually monitored, still there are traces of anomalies such as entry of an unauthorized person inside the area of video surveillance, the person hiding objects within his body, carry bags etc.. Some objects might be an unauthorized as per the norms of the area under video surveillance. The objective of the manuscript is to make the system intelligent to detect completely occluded objects to fill the lapse in security. As per the guidelines specified by the ministry of Urban Development, Government of India (http://smartcities.gov.in/writereaddata/smartcityguidelines.pdf) video crime monitoring is one of the major features that a smart building should carry under E-Governance and Citizen Service. To support video crime monitoring several advancements are made in building a video analytics starting from face detection, face recognition, object detection, object identification, object tracking, car detection, animal and non- animal identification etc.. as discussed in related works below, but the problem of handling hidden object detection under video crime monitoring still remains an unsolved problem in research. Even the human intelligence with the eye of resolution 576 megapixels fail often in detected the completely occluded objects. Edgar Osuna, et al [1] used Support Vector Machines (SVMs) to detect a face from non- face. The system proposed detects face image from front view. A decision boundary is created to differentiate face from non- face. The system detects only frontal faces in the grayscale image since only frontal face images are trained. The detection rate of the system is 97.1%. Jeffery Huang [2] used decision trees to separate face and non-face images. They used 2,340 face images from the FERET database for discriminating face and non-face images. The rules used for the proposed approach detect facial features. Threshold measure is used to find the presence of a face. The presence of a face

is confirmed when the values obtained exceeds the threshold. The accuracy of the face detection system is claimed to be 96%. L. Sirovich et al [3] came up with face recognition algorithm that recognizes the person whose face is pre-trained. The database contains 2500 images and sixteen persons face is used to capture the images. The recognition rate is claimed to be 85% when tested with different lighting conditions. Detection and Recognition of objects that included face - nonface[4], living being- non-living being, recognizing the trained objects [5] is a challenging approach in the sectors of robotic vision. The further challenge involved is to track in the frame by frame approach [6]. The success lies in localization and extraction of the particular region of interest in the image or video. Pentland et al [7] used PCA to reduce the dimension of the feature vector and the images are trained with the concept of Eigen faces for face recognition. The experimental analysis says that the system has a recognition rate of 95%. Aili Wang [8] used Wavelet transform and improved 2DPCA to extract feature and the classifier used to for training face image is nearest neighborhood classifier. ORL dataset is used for experimentation. The recognition rate of the system is claimed to be 92% when ORL dataset is used for experimentation.

## II. WORKFLOW AND METHOD DESCRIPTION

. The system needs both Internet Protocol Camera and Thermal camera, which is mounted in the entry of the building as shown in Figure 1

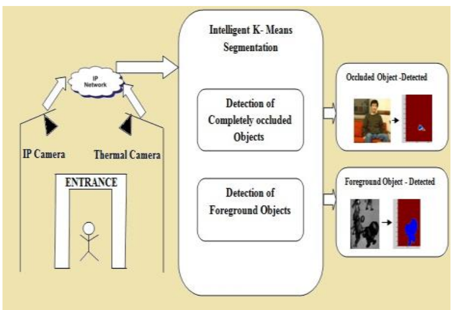



**Fig 1.** Architecture diagram of the proposed approach

The resultant image obtained after adding RGB image and Infra-Red image is named as Image A for convenience as shown in figure 2. While linearly blending two images the images should be exactly similar in a total number of horizontal and vertical pixels. The image A is smoothened by Gaussian filter and the resulting image is named as image G for convenience. Then the image G is again added to the IR image, and the resulting image obtained is named as image B for convenience.

The image B is taken as input and the back-ground of the image B is eliminated and the back-ground subtracted image of image B is named as Image C. Then the RGB image is taken and it is converted into HSV image and it is named as image HSV for convenience. The image C (Background subtracted image) and image B is added together to get a fused image, where we can see the object hidden. Now the task is to segment and detect the object hidden. The fusion image is converted into a grayscale image and the proposed intelligent K-means segmentation algorithm is used to segment the objects hidden and then the object is detected with the bounding box.

K-means clustering algorithm is used to segment the occluded object. K-Means Clustering algorithm comes under unsupervised learning since we don't have any training data with prior knowledge about its class to train them and classify them.

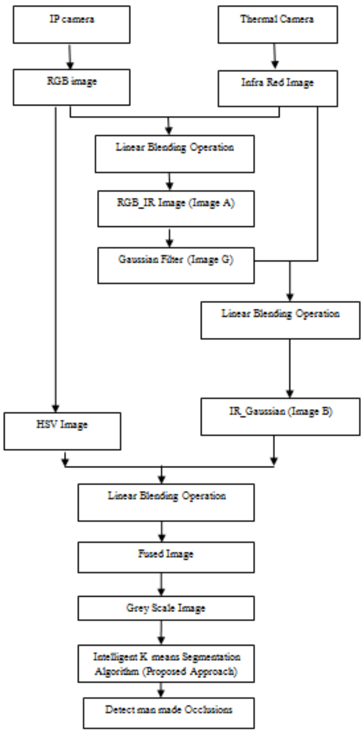



**Fig 2.** Work Flow of the Proposed Approach

In K- Means algorithm we have to do knowledge discovery process based on distance measure and cluster the data. Since there are chances for the presence of n number of objects of varying size, color, texture etc… the detection of objects in the area under video surveillance is an unsupervised learning task. The success of the segmentation depends on the value of K. hence the problem behind is to choose an optimal value for K because choosing K by trial and error method. The segmentation algorithm has foreground detector which To overcome this problem the proposed algorithm finds an optimal value for K as discussed below.

The minimum value of k =1 hence there is no clustering operation all the features are grouped into a single cluster hence the minimum value ok k be 2 because based on the assumption of child and adult the features are clustered. The maximum value of k depends on upon the number of feature vectors. Let n be the number of feature vectors and the maximum value of k is chosen to be n-1.

## III. Mathematical Model for Segmentation

To for clear elaboration let us take a 5*4 window from the Red Green Blue image that is given as input.

**Table 1:** Representing 5*4 image segment with its corresponding RGB Values

| R | G | B |
|---|---|---|
| 11 15 67 89 | 21 25 77 99 | 33 37 89 111 |
| 6 64 4 56 | 16 14 66 74 | 28 26 78 86 |
| 6 48 7 44 | 15 16 53 57 | 30 31 68 72 |
| 9 46 7 47 | 18 16 56 55 | 33 31 71 70 |
| 8 4 42 47 | 179 183 242 226 | 25 38 60 79 |

The Hue Saturation Value color space has different parameters for the HUE, the saturation and the value (intensity) of the pixel, and so we get easy access to the actual Hue (variation in color) of the pixel, without any concerns about different intensity levels. Hence the value of k depends on upon the color variations in the RGB image and thus the variations in color are captured in hue. For example, in the above hue value there are two variations in color thus the upcoming similar pixels are calculated for minimum distance with respect to the centroid 1 and centroid 2 and then the classes are labeled.

**Table 3.** Grouping pixels Hue with their corresponding centroids based on similarity

| Centroid 1 (0.333333333333333) | Centroid 2 (0.305555555555556) |
|---|---|
| 0.333333333333333 | 0.305555555555556 |
| 0.333333333333333 | 0.305555555555556 |
| 0.333333333333333 | 0.305555555555556 |
| 0.333333333333333 | 0.305555555555556 |
| 0.333333333333333 | 0.305555555555556 |
| 0.333333333333336 | 0.305555555555556 |
| 0.333333333333336 | 0.305555555555556 |
| 0.333333333333336 | 0.305555555555556 |
| | 0.305555555555556 |
| | 0.305555555555556 |
| | 0.305555555555556 |
| | 0.305555555555556 |

## IV. Choosing Optimal Value for K

More centroid thrown will saturate the space with clusters of minimum distance. So that's not a very good solution. The total number of cluster is plotted in X axis and the variance (aggregate distance from the data point to the cluster) is plotted in y-axis as shown in the Figure. The key observation to be noted and updated while choosing the k value is to measure how much the variance change as one more cluster is added? At each step increment, the cluster count or k value and measure the variance that is the intra-cluster distance should be minimized when cluster count gets incriminated. Increment the cluster count in the x-axis and find the variance observed for each K value incremented. The variance or intra-cluster distance is the average between the point x and the nearest centroid cj as shown in the equation. Logic: Check at from which point (cluster count or K) in the graph there is a huge change in variance. The distance measure used for computation is Euclidean distance. Since Euclidean distance is based on the minimal distance between the data points it is chosen to measure similarity in the proposed approach.

**Fig 4.** RGB image captured

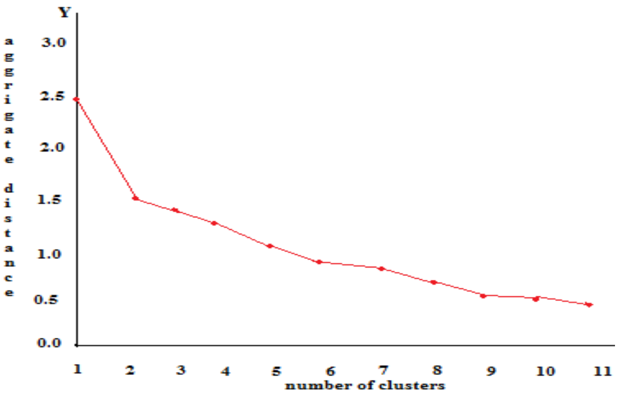



**Fig 3.** Choosing optimal value of K using the proposed approach

In the existing system, the value of k is determined by trial and error method on seeing the segmentation results. But the proposed algorithm has the intelligence to find the optimal value of k based on the logic given below. For every value of K, the variance is calculated. K value is finalized in such a way that the magnitude of variance for the current K value and the previous K value should be larger. For the given input image the abrupt change in variance is observed when k =2. The variance for K =1 is 1.5 and the variance for K =2 is 2.5 hence the change in variance value is found to be 1 in this case. When a change in variance is observed for K value above 3, it is found to be lesser that one and it decreases constantly as shown in the figure. In short, we can say that k value is chosen on seeing that how much the variance change when one more cluster is added. For the given image as input, the K value is chosen as 2 as per the proposed approach.

V. EXPERIMENTAL RESULTS

The figure shown below is the RGB image obtained from IP camera and figure 4 is the IR image of the same person captured in thermal camera.

**Fig 5.** IR image

**Fig 6.** Addition of RGB and IR image (image A)

**Fig 7.** Image A after Gaussian blur

**Fig 8.** (Image B) Addition of IR Image and Gaussian blur image

**Fig 9.** HSV Image

**Fig 10.** After adding HSV and Image B

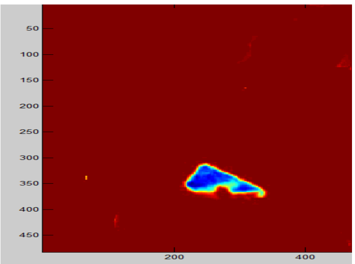



**Fig 11.** Segmented Object

Table 5 Performance Evaluation of the proposed approach using FLIR ONE images (as shown in Table 4)

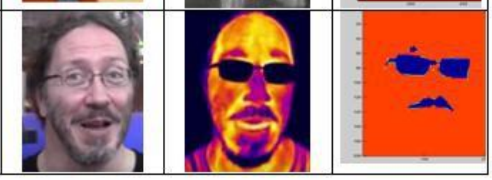
| S.NO | RGB Image / IR Image / Segmented Objects |
|---|---|
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 |  |

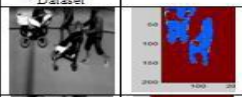
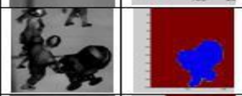
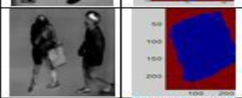
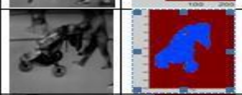
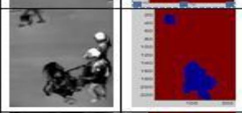
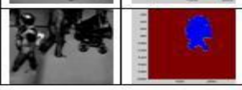
**Table 4:** Segmented Image from the foreground

| S.No | Total Number of Objects in the image | Number of Segmented Objects (True Positive) | Objects Falsely Segmented (False Positive) | Precision (Segmentation Rate) |
|---|---|---|---|---|
| 1 | 8 | 6 | 0 | 75 |
| 2 | 1 | 1 | 0 | 100 |
| 3 | 2 | 2 | 0 | 100 |
| 4 | 1 | 1 | 0 | 100 |
| 5 | 1 | 3 | 2 | 30 |

Precision = (true positive / true positive + false positive) *100 = 81% False Discovery Rate= 100- precision = 19%

recorded while capturing the data is 20-degree celsius. The performance of the proposed system is tested using Stereo Thermal Dataset and the precision is being calculated as shown in table 6

**Table 6.** Evaluation of Performance using Stereo Thermal Dataset

| S.No | Reference Image Frame from Stereo Thermal Dataset / Segmented Objects | Total Number of Objects in the image | Number of Segmented Objects | Precision |
|---|---|---|---|---|
| 1 |  | 2 | 2 | 100 |
| 2 |  | 1 | 1 | 100 |
| 3 |  | 1 | 1 | 100 |
| 4 |  | 1 | 1 | 100 |
| 5 |  | 1 | 1 | 100 |
| 6 |  | 3 | 1 | 33.33 |

Precision = (true positive / true positive + false positive) *100 = 88.89%

False Discovery Rate= 100- precision = 11.11%

**Performance Evaluation**

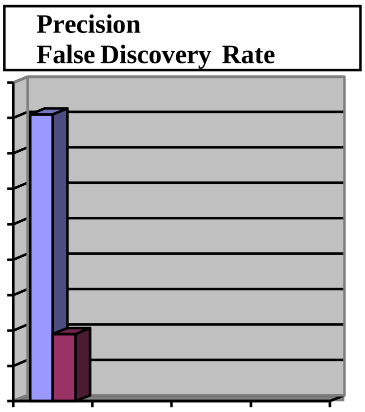



**Fig 13.** Performance Evaluation of the proposed approach using Image frames of Stereo Thermal Dataset

## VI. CONCLUSION

The intelligent K-means segmentation algorithm segments the objects present in the foreground and the completely occluded objects by choosing the optimal value for K. The proposed approach holds a huge scope in open access buildings such as airports, railway station, malls etc.. and the buildings that need to be highly secured. The proposed algorithm is tested with images collected from FLIR ONE and Stereo Thermal Dataset and the system achieves 81% and 88.9% of precision respectively.

## VII. REFERENCES:

[1]. Osuna, Edgar, Robert Freund and Federico Girosit, Training support vector machines: an application to face detection. Computer vision and pattern Recognition pp. 130-136 ( 1997).

[2]. Huang, Jeffrey, Srinivas Gutta and H. Wechsler. Detection of human faces using decision trees. Automatic Face and Gesture Recognition, Proceedings of the Second International Conference on IEEE Pp. 248-252 (1996).

[3]. Kirby, Michael and Lawrence Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Transactions on Pattern analysis and Machine Intelligence 12(1): 103-108 (1990).

[4]. K.S. Gautam and T.S. Kumar, Discrimination and Detection of Face and Non-face Using Multilayer Feedforward Perceptron, *In Advances in Intelligent Systems and Computing published* – Springer – *from the proceedings of International Conference ICSCS-15*. 397: 89-103 (2016).

[5]. Senthil Kumar, T, Saivenkateswaran, Evaluation of Video Analytics for face detection and recognition. International Journal of Applied Engineering Research 10(9): 24003- ( 2016).

[6]. T. Senthil Kumar, K.S. Gautam, H. Haritha, Debris Detection and Tracking System in Water Bodies Using Motion Estimation Technique, In Advances in Intelligent Systems and Computing published – Springer - from the proceedings of the 6th International Conference on Innovations in Bio-Inspired Computing and Applications 424: 275-284 (2015),

[7]. Turk, Matthew A. and Alex P. Pentland. Face recognition using eigenfaces." Computer Vision and Pattern Recognition, Proceedings CVPR'91., IEEE, Pp. 586-591 (1991).

[8]. Wang, Aili, Na Jiang, and Yuan Feng, Face Recognition Based on Wavelet Transform and Improved 2DPCA. Instrumentation and Measurement, Computer, Communication and Control-IMCCC Pp. 616-619 (2014)

[9]. T. Senthil Kumar and G. Narmatha, Video Analysis for Malpractice Detection in Classroom Examination, In Advances in Intelligent Systems and Computing published–Springer- from the proceedings of International Conference ICSCS 2015, 397: 135-146 (2016) ISBN: 978-81-322 -2669-7 (Print) 978-81-322-2671-0

[10]. T. Senthil Kumar and G. Narmatha, Suspicious human activity detection in classroom examination, In Advances in Intelligent Systems and Computing-Springer from the proceedings of International Conference on Computational Intelligence, Cyber Security and Computational Models, ICC3-15, 412: 99-108 (2016).

[11]. Li, L., Gong, R. and Chen, W., Gray level image thresholding based on fisher linear projection of two-dimensional histogram. *Pattern Recognition 30* (5): 743-749 (1997).

[12]. Abo-Eleneen, Z.A., Thresholding based on Fisher linear discriminant. *Journal of Pattern Recognition Research* 2: 326-334 (2011)

[13]. Sharma, L., Yadav, D.K. and Singh, A., Fisher's linear discriminant ratio based threshold for moving human detection in thermal video. *Infrared Physics & Technology 78:* 118-128 ( 2016)