

# A Hybrid Approach for Image Segmentation in the IoT Era



Tallha Akram, Syed Rameez Naqvi, Sajjad Ali Haider,  
and Nadia Nawaz Qadri

## 1 Introduction

The Internet of Things (IoT) has a potential to transform our world by intelligently connecting every device that touches both of our professional and everyday lives through a range of sensors, cloud computing, etc. and also has a capacity to unlock the door to the next industrial revolution. In the last decade, IoT has covered numerous domains including communication [1, 2], energy efficiency [3, 4], vehicular systems [5], smart farming [6], healthcare [7], computer vision [8], and to name a few. Detecting object of interest from an IoT-based camera device can be useful, especially under the scenario of identifying one dedicated class from a bulk of processed information.

Spectral clustering is a class of graph theoretic procedure, which is popular for finding natural groupings. Over the last decade, it has become a widely adopted tool for image segmentation problems via the normalized cut (NCut) criterion [9]. Spectral clustering popularity increased substantially among the researchers in the field of computer vision, bioinformatics, robotics, etc. The idea has proven itself to be an active area of research among machine learning community. The core reasons behind are its capability to accurately group complex structured data and its simplicity in implementation. Image segmentation using spectral clustering is not an expedient way of dealing large size images due to its high computational demand and memory requirements. To overcome such predicament, a dimensional reduction procedure is integrated as an effective preprocessor before calculating an affinity matrix – which combats the curse of dimensionality. In the situation where computational bottleneck is comprehended, a competent selection of suitable

---

T. Akram (✉) · S. R. Naqvi · S. A. Haider · N. N. Qadri  
Department of Electrical and Computer Engineering, COMSATS University, Islamabad, Pakistan  
e-mail: [tallha@ciitwah.edu.pk](mailto:tallha@ciitwah.edu.pk)



**Fig. 1** A pictorial illustration of weighted binary tree-based fast spectral clustering

preprocessor reduces the bottleneck risk by reducing the data structure size. Selection of preprocessor is a key task and has foremost effects on clustering.

In this work, we propose an unsupervised image segmentation technique using spectral clustering aimed at salient object extraction. Few challenges faced by image segmentation based on spectral clustering are its infeasibility to process large images due to high computational demand, memory requirements, and its sensitivity to irrelevant and noisy data. These challenges are successfully addressed in the proposed work of weighted binary tree-based fast spectral clustering (WBTFSC); sample results are demonstrated in Fig. 1. The algorithm integrates dimensionality reduction method into spectral clustering by introducing an effective preprocessor, which is comprised of two fundamental steps: (1) color quantization based on weighted binary tree and (2) unique pixels selection. Noisy images are also tested with proposed algorithm and showed exceptional results comparable to the other implemented techniques.

Many clustering techniques are strictly bound to Euclidean geometry, making assumption explicitly or implicitly to form convex regions in Euclidean space, compared to be more flexible and confined to wider geometry range of spectral clustering [10].

Color quantization is deemed to be a prerequisite of many color image segmentation algorithms, and we exploited this key to be an effective preprocessor [11–13]. The contribution is twofold: Firstly, we leveraged the fact that natural images include fewer consistent groups than pixels [14] and introduced an effective preprocessor of weighted binary partition tree for quantization. Secondly, selection of unique pixels from normalized image further reduces the data structure size before subjecting to spectral clustering. Labeling is finalized using fuzzy C-means soft clustering to exhibit fuzzy behavior and making it feasible for salient object extraction with fused clusters.

In spectral clustering, image is represented by a weighted graph, where each pixel is a node connected to other nodes with edges. Despite the fact, spectral clustering achieves accurate clustering and makes precise segments, it is still not considered to be a strong competitor to other clustering techniques [15, 16]. Spectral algorithm forms affinity matrix of size  $(n \times n)$ , where  $n$  is the total number of pixels in the image, and computes eigenvectors of this affinity matrix with computational complexity of  $O(n^3)$  and memory requirement of  $O(n^2)$ .

To improve the performance of spectral clustering, multiple schemes were proposed and implemented [17, 18]. One technique is the providence of original data representatives, which should have the capacity of retaining original values. Exploiting the subsampling of randomly selected data w.r.t. to grouping criteria is another reasonable approach. Low-rank approximation in spectral clustering is another central option for the dimensionality reduction. High-quality clustering achieved by spectral clustering can be annihilated with poor choice of preprocessor, which must be selected carefully for improved performance.

## 2 Related Work

Spectral clustering for image segmentation is a graph theory-based information extraction procedure which describes the image as a weighted graph and partitions them using optimized cost function. Segmentation is done under the potency of feature vectors (directly/indirectly), based on eigendecomposition of graph Laplacian matrix [19]. Possibly the first comprehensive effort was made by Shi and Malik [20] to segment image using spectral clustering. They formulated the problem from a graph theoretic perspective and introduced a normalized cut to segment the image.

Traditionally spectral clustering performance degraded hastily when dealing with high-resolution images due to its high computational cost. To overcome this problem, few techniques are implemented depending on the nature of the problems. Bruce et al. [18] used the multilevel approach for graph partitioning in which sequence of increasingly smaller graphs approximate a complete graph. The smallest graph is then partitioned using spectral clustering and finally back propagated to graph hierarchy. Same spirit of multilevel approach is discussed by Yan et al. [21] where coarsening and un-coarsening methodology are followed and algorithm is designed to deal with the computational requirements to overcome spatial bottleneck. They employed fast approximate spectral clustering using K-means clustering and random projection trees as a preprocessing step. The intension was to minimize data reduction effect on clustering accuracy, in the spirit of rate-distortion theory.

Another approach was put forth by Ducoumau et al. [22], where clustering was configured as hypergraph cut problem to optimize the objective function. The proposed method was based on the principle of multilevel paradigm with three key steps, starting with the initial step of hypergraph reduction, followed by a spectral

clustering of reduced hypergraph, and the final step of cluster refinement. Similarly, Tung et al. [23] implemented a new distinctive approach of block-wise processing in combination with stochastic ensemble consensus. They first divided the image into multiple nonoverlapping fix-sized blocks and then calculated affinity matrix for each block. Finally, they merged the segments using stochastic ensemble consensus.

Tasdemir et al. [26] proposed approximate spectral clustering in combination with neural networks by employing self-organizing maps and neural gas as quantizer – a preprocessing step. They exploited local density-based similarity measure without user-predefined parameters. Similarly, a fast affinity propagation clustering approach was implemented by Shang et al. [20] in which they considered both local and global structural information. The concept was based on multilevel graph partitioning which managed to implement both vector-based and graph-based clustering.

Fowlkes et al. [14] made an exertion to overcome the time and computational limitations by incorporating the low-rank approximation. They introduced Nystrom-based spectral clustering algorithm for image segmentation. The main idea was to solve the grouping problems for small random subsets of pixels in image and then anticipate this solution to all pixels. To improve Nystrom spectral clustering, few methods were also proposed in literature by applying different quantization schemes [27]. All techniques referred above are based on the concept of quantization before object of interest detection. A concept of finding salient regions is also presented which is based on attention mechanism and considers contrast to be a primary factor.



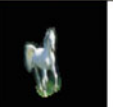
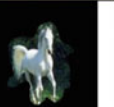

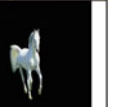













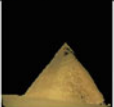
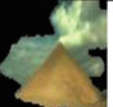

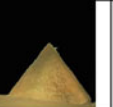




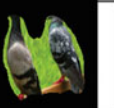
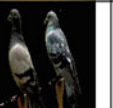





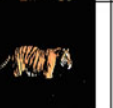









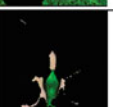



Goferman et al. [28] proposed context-aware saliency detection. The objective is to detect the salient region(s) with respect to a scene. They considered both low level and global feature. Cheng et al. [29] proposed histogram-based contrast method to measure saliency. One of the biggest challenges in saliency detection is to acquire salient regions w.r.t. scene understanding.

### 3 Clustering Methods

The proposed design is a combination of multiple clustering methods including hierarchical clustering (weighted binary partitioned tree), graph clustering (spectral), and finally fuzzy clustering (C-means). Sample results of proposed technique are demonstrated in Fig. 2.

#### 3.1 *Weighted Binary Partition Tree*

Let us consider an image  $S$  of size  $(n \times n)$  for  $(m = n)$  where  $s = (r, c)$  for  $(s \in S)$  to be row and column indices. Initialize a binary tree clustering by partitioning  $S$  into

-	Original Image	Spectral Segmentation	CEV	GBVS	Proposed Method	Ideal Segmentation
a						
b						
c						
d						
e						
f						
g						
h						

**Fig. 2** Visual comparisons with respect to object of interest detection; (1) original image, (2) spectral algorithm for object extraction, [1] (3) curve evolution algorithm using low depth field, [17] (4) graph-based visual saliency [18], (5) WBTFSC algorithm for object extraction, (6) ground truth

$K^1$  disjoint sets, subject to the constraint of achieving maximum of two nodes (*left and right*), where the numbers of clusters are already defined.

Each partitioned node belongs to its parent node which flows up in the hierarchy to the root node “X” – the starting point of any partition. Let  $P_n$  represents set of image pixels which correspond to node  $n$ . For the proposed weighted binary tree, let the head node ( $n$ ) be represented by a number “1,” while the children nodes are  $2n$  and  $2n + 1$ . The partition of  $X$  is on the basis of colors for the selected image. The partition is controlled by cluster colors which rely on mean and variance value and

stored as linked list. In splitting process node, order is very important; following the procedure in an order, we need to state the second-order statistical properties.

$$R_n = \sum_{s \in P_n} W_s x_s x_s^t \quad (1)$$

$$O_n = \sum_{s \in P_n} W_s x_s \quad (2)$$

where minimum squared deviation is given by a cluster mean; therefore, it favors an assumption that mean is equal to  $q$  – the cluster's quantization value.

$$\mu_n = O_n / |P_n| \quad (3)$$

Cluster covariance is defined as:

$$\sum_{n+1} = \sum_n - \frac{1}{|P_n|} O_n O_n^t \quad (4)$$

Weights are assigned according to relation:

$$W_s = \left( \frac{1}{k_\sigma * \min(\|\nabla S_{xy}\|, 14 + 2)} \right)^2 \quad (5)$$

where  $K_\sigma$  is a Gaussian smoothing kernel which filters gradient estimator and  $S_{xy}$  is color range with minimum values of 2, the lowest visually indistinguishable level.

This approach is modified to show improved performance for large clusters, which endorses the maximum cluster's variation by handling and splitting order. Unit vector  $\hat{u}$  is defined as:

$$\sum_{s \in P_n} ((x_s - \mu_s)^t \hat{u})^2 = \hat{u}^t \sum_{n+1} \hat{u} \quad (6)$$

The eigenvector  $\hat{u}$  corresponds to the largest eigenvector  $\lambda_n$  of  $\sum_{n+1}$ . Total squared variation in direction of  $\hat{u}$  is given as:

$$\sum_{s \in P_n} ((x_s - \mu_s)^t \hat{u})^2 = \lambda_n \quad (7)$$

The points in the  $Wp_n$  are divided into  $Wp_{2n}$  sorted as  $Wp_{2n+1}$

$$Wp_{2n} = \left\{ s \in P_n : \hat{u}_n^t x_s \leq \hat{u}_n^t \mu_n \right\}$$

and  $Wp_{2n+1}$  is defined as:

$$Wp_{2n+1} = \left\{ s \in P_n : \hat{u}_n^t x_s > \hat{u}_n^t \mu_n \right\} \tag{8}$$

New weighted statistics are created as:

$$R_{2n} = w_n R_n - w_{2n} R_{2n}$$

$$O_{2n+1} = w_n O_n - w_{2n} O_{2n} \tag{9}$$

$$P_{2n+1} = |w_n P_n| - |w_{2n} P_{2n}|$$

where the weights  $w_n$  and  $w_{2n}$  are static to be 0.6 and 0.4, respectively, for optimal performance. The optimal splitting order is selected with the objective of maximum reduction in error squared, defined as:

$$\text{TSE} = \sum_{\text{All Leaves } n} \sum_{s \in p_n} \|x_s - \mu_n\| \tag{10}$$

Here we utilized the largest principle eigenvector to achieve the best split (Fig. 3).

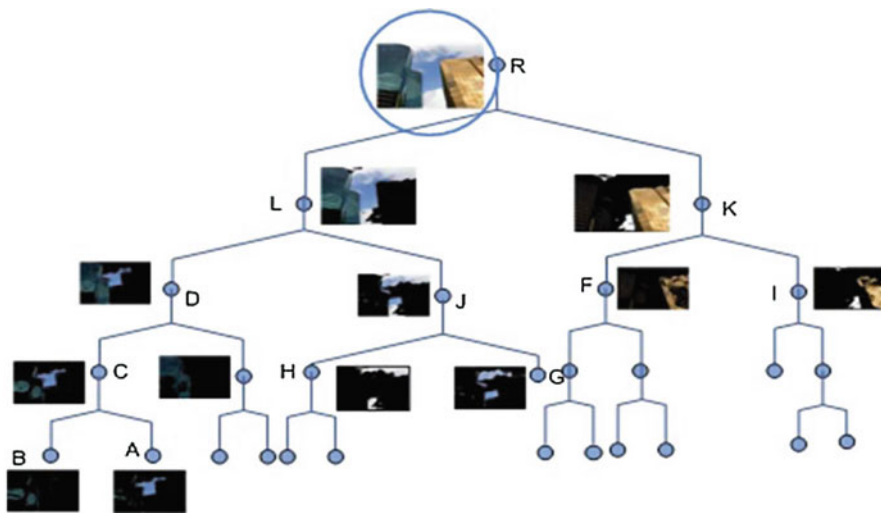


Fig. 3 Cluster partitioning using weighted binary partition tree

### 3.2 Spectral Clustering

Spectral clustering for image segmentation is a graph theory-based information extraction procedure which describes the image as weighted graphs and partitions them using optimized cost function. Let  $G(V, E)$  be an undirected graph, where  $V = (v_1, v_2, v_3, \dots, v_n)$  are a set of  $n$  nonempty vertices along with set of edges,  $E$ . The graph is weighted; each edge between vertices  $v_i$  and  $v_j$  have a weight of  $w_{ij} > 0$ .  $G'$  is subgraph of  $G$  having vertices and edges  $(V', E')$  with the fact  $V' \in V$  &  $E' \in E$ . The weight adjacency is given as  $W = (w_{ij})_{n \times n}$ , where  $w_{ij} \in \Re^{(n \times n)}$ . For weight values  $(w_{ij}) = 0$ , correspond to the fact that vertices  $V_i$  and  $V_j$  are not connected.

Different techniques like  $K$ -neighbors,  $k$ -nearest neighbors, and fully connected graph are used to transform given data points  $(x_1, \dots, x_n)$  into pairwise distance  $d_{ij}$  or pairwise similarities  $s_{ij}$  into graph. Let  $d_{ij}$  be the dissimilarity matrix; the weight matrix  $W = (w_{ij})$  for  $(i, j) = 1, \dots, n$  can be calculated from sigmoid function as:

$$w_{ij} = \exp -\|x_i - x_j\|^2 / 2\sigma^2 \text{ For } i \neq j \text{ \& } 0 \text{ Otherwise.} \quad (11)$$

The parameter  $\sigma$  has supplementary impact on clustering obtained. The selection of  $\sigma$  is a very critical step; one approach is to run an algorithm for different values of  $\sigma$  until the least squared intra-cluster distance to its center position is obtained [23]. Mentioned hit and trial rule for optimal value of  $\sigma$  looks not feasible for large datasets, so few other strategies are instigated [31]. We selected sigma value 1 for the optimized solution.

The degree matrix  $D$  is defined as diagonal matrix with  $D = [d_{ij}]_{n \times n}$ , where  $d_{ij} \in \Re^{n \times n}$  defined as  $d_i = \sum_{j=1}^n w_{ij}$ . un-normalized graph Laplacian is given as  $L = D - W$ , which may vary in case of unweighted graph as  $L = D - A$ , where  $A$  is adjacency binary matrix of  $G$  given as  $A = a_{ij}(n \times n)$ , where  $a_{ij} = 1$  if edge connection is found between  $V_i$  and  $V_j$  and 0 if otherwise. The Laplacian matrix must satisfy the following properties. For vector  $X \in \Re^{n \times n}$ :

$$x^T L x = \frac{1}{2} \sum_{i,j=1}^n w_{ij} (x_i - x_j)^2 \quad (12)$$

$L$  is symmetric positive definite. The smallest eigenvalue of  $L$  is 0, with corresponding eigen indicator vector 1, i.e.,  $1 = (1, 1, \dots, 1)^T$ . Luxburg [32] theoretically proves all the properties of graph Laplacian matrix. Fiedler suggested that  $\lambda_2$  is the second eigenvector of  $L$  that represents relationship among connected graph components also known as fielder's eigenvector. The normalized Laplacian matrix is given as  $L = I - D^{-(0.5)} W D^{-(0.5)}$ . Second random-walk Laplacian matrix is given as  $L_{rw} = I - D^{-1} W$ .



## 4 Proposed Methodology

The proposed object extraction algorithm is summarized with flowchart shown in Fig. 4, which incorporates three fundamental stages; the first stage deals with color quantization using weighted binary tree, where the first loop divides the image into distinct clusters, while the second loop

assigns each cluster a unique color value from the palette. The second stage deals with the normalization procedure and finding unique pixels in the quantized image and followed by final stage of object of interest detection and later extraction using spectral clustering.

Algorithm analysis revealed us that reduction of a color palette size with weighted binary partition tree quantization still is not very effective in calculating reduced similarity matrix of size  $(n \times n)$ , considerably when image is of large size. The reduction of distinct colors as an outcome of binary tree quantization has no effect on total number of nodes subjected to spectral clustering but to the number of natural colors. To overcome such predicament, all pixels are normalized in the range of  $(0 \rightarrow 1)$  before selecting unique values to construct similarity graph. Essentially, duplication removal process keeps in record total number of duplicate pixels and their respective addresses for each basic RGB value. This procedure allows us to deal with large size images in calculating affinity matrix, which is one of the prime steps in spectral clustering.

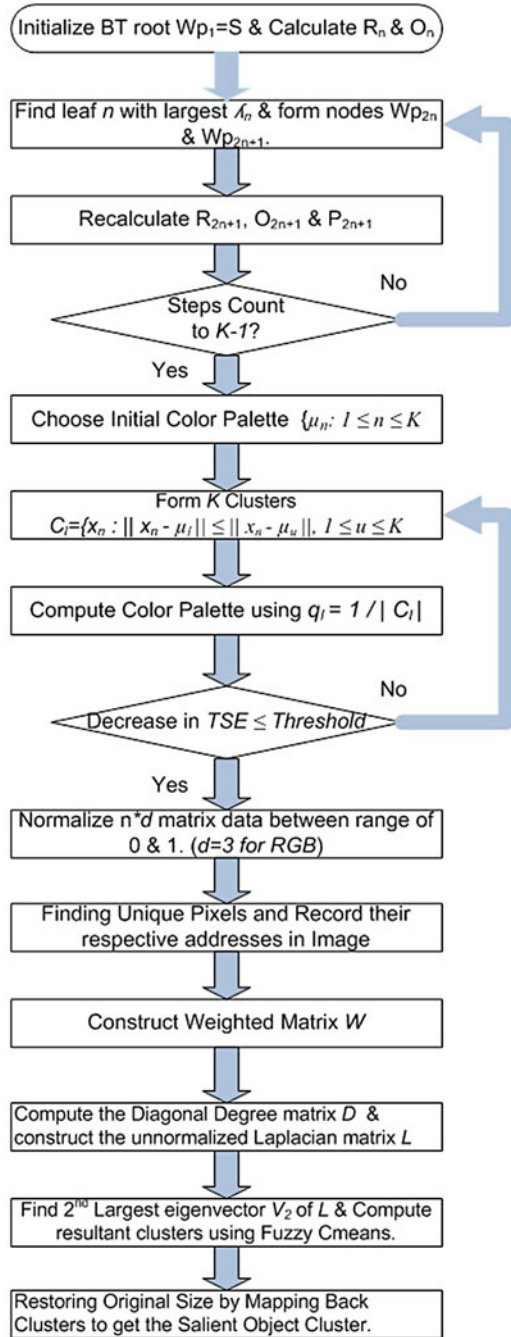
The number of clusters  $K$  is defined by user under constraint that the weighted binary tree quantization clusters are greater than spectral clustering, so not to annihilate the visual contents of image but to reduce the number of distinct colors. If number of clusters created using spectral clustering is insufficient in number, it will result in fused segments of object and background, resulting in false segmentation.

For similarity matrix, sigmoid function approach is used instead of  $K$ -nearest neighbors and  $\epsilon$ -neighborhood graph, due to large number of connected components. Ultimately, grouping is done using fuzzy C-means algorithm, and cluster members are mapped to their respective addresses with the duplication sequence. The method shows performance degradation if preprocessor is not defined well or image having uniformly distributed pixels.

Based on the evaluation of WBTFS in obtaining meaningful segmentation for object of interest detection, spectral clustering fulfills two conditions:

- Nodes that represent possibly larger connected regions belong to object of interest class and should be selected.
- Regions represented by selected clusters should be disjoint.

**Fig. 4** General flow diagram of weighted binary tree-based fast spectral clustering



## 5 Results and Discussion

To assess the performance of proposed algorithm from the perspective of object of interest detection and extraction, a set of images are selected from MSRA<sup>1</sup> salient object database, Berkeley<sup>2</sup> segmentation dataset, and image collection of our group, with the perspective of finding at least one obvious salient object. To show the consistency and effectiveness of our proposed algorithm, performance measure was credited in respect of accuracy and efficiency. Results are visually compared with three state-of-the-art algorithms. Analysis is made from the perspective of object of interest extraction. In order to show the consistency and effectiveness of the proposed WBTFSC algorithm, different performance measures are credited, in respect to time, size, and efficiency. To demonstrate the performance of our proposed algorithm, we compared our results with three state-of-the-art algorithms, spectral clustering [30], curve evolution with low depth of field [24], and graph-based visual saliency (GBVS) [25].

Few images are selected for visual comparison between proposed technique and selected methods as shown in Figs. 2, 5, and 6. Ideal segmentation or ground truth is shown in column 6 which are manually segmented images to calculate accuracy. Our approach has shown notable results, visually observed in column 5, when compared with conventional spectral clustering [31] in column 2, curve evolution approach in column 3, and graph-based visual saliency in column 4. It is observed that not only the test images with single salient object showed good performance but test images with multiple salient objects are also segmented and extracted successfully and can be observed in images “b,” “e,” “I,” “n,” etc.

With weighted binary tree palette design, few image pixels are misclassified and covered with wrong cluster labels which results in the loss of object information with spectral clustering object extraction procedure. In order to coup with this situation, we took advantage of soft clustering of fuzzy C-means algorithm. It assigns data member  $x_i$  not to strictly bind to one cluster, rather deal it in flexible way. For the proposed technique, object extraction principle is based on color distinctness where each color represents different cluster. The presence of multiple clusters within salient object results in the loss of inner details of object and can be observed in images “f,” “k,” and “p,” but objects’ contours remain preserved.



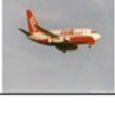
Table 1 shows time comparison of all selected techniques and size comparison of proposed method and spectral clustering, which can be analyzed with respect to similarity graph time and total processing time to extract an object.

It can be noticed that WBTFSC has shown a remarkable reduction in similarity graph time to create  $(n \times n)$  matrix. The maximum time for spectral method is 45.27 s for image “e” as compared to 9.28 s for WBTFSC. Minimum time for

---

<sup>1</sup>[www.research.microsoft.com/en-us/um/people/jiansun/SalientObject/salient\\_object.htm](http://www.research.microsoft.com/en-us/um/people/jiansun/SalientObject/salient_object.htm)

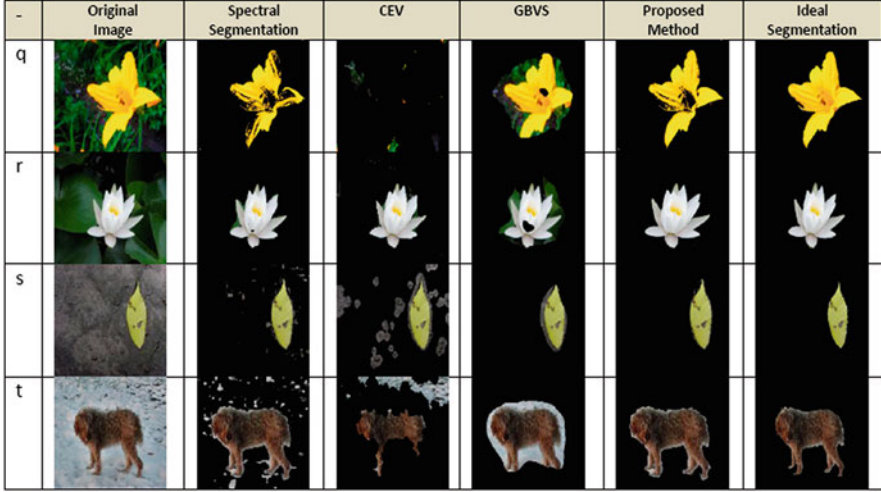
<sup>2</sup>[www.elib.cs.berkeley.edu/segmentation](http://www.elib.cs.berkeley.edu/segmentation)

-	Original Image	Spectral Segmentation	CEV	GBVS	Proposed Method	Ideal Segmentation
i						
j						
k						
l						
m						
n						
o						
p						

**Fig. 5** Few more visual comparisons with respect to salient object extraction. (1) Original image; (2) spectral algorithm for object extraction [1]; (3) curve evolution algorithm using low depth field [17]; (4) graph-based visual saliency [18]; (5) WBTFSC algorithm for object extraction; (6) ground truth

WBTFSC is 1.30 s for “p” image compared to spectral clustering, 19.17 s and curve evolution approach, 34.31 s.

GBVS has shown minimum time for object detection, but results are in contrary to our approach in terms of accuracy. A comparison of proposed method and spectral clustering in terms of execution time and similarity graph size is given in Fig. 7. Notable difference can be observed in execution time and similarity graph size where WBTFSC maximum nodes are approximately 16 K in size and execution time less than 20 s in contrast to spectral clustering.



**Fig. 6** Few more visual comparisons with respect to salient object extraction. (1) Original image; (2) spectral algorithm for object extraction; (3) curve evolution algorithm using low depth field; (4) graph-based visual saliency; (5) WBTFSC algorithm for object extraction; (6) ground truth

Spectral clustering performance degraded hastily when showed up with noise. This deficiency is challenged with our proposed method and showed exceptional results, demonstrated in Fig. 8. We introduced three different types of noises with different intensity levels and extracted the object of interest with good accuracy. It can be observed that our proposed technique showed better results especially when salt-and-pepper noise is introduced. Table 2 shows similarity graph time and total time for noisy images. For each column assigned to different noises, left sub-column shows our proposed method processing time, and right sub-column shows spectral clustering’s processing time.

In order to objectively evaluate the segmentation performance from the perspective of object of interest detection and extraction, we consider two different strategies. For the first, output image is segmented into regions  $(R_1, R_2, \dots, R_n)$ .  $R_i \cap R_j = \phi$  for  $i \neq j$  and  $\cup_{i=1}^n R_i = I$ . For each segment  $R_i$  in image, if more than 50% overlap with ground truth foreground  $O_{GT}$ , it counts as foreground  $R$  [34–35].

We tested the efficiency of images by dividing into 164 regions/segments; comparative plot is shown in Fig. 9 (right).

$$R_{obj} = \cup_{(i:(p_{R_i, O_{MS}} > 0.5))} R_i. \tag{13}$$

or

$$p(R_i, O_{MS} \geq p(R_k, O_{MS}), \forall k \in [1, n_s]) \tag{14}$$

**Table 1** Time and size parameters for the proposed technique of WBTFSC in comparison with curve evolution method, GBVS, and spectral clustering

Figure#	Spectral method		CEV		CBVS		BTFC		Q-B-tree	S. graph size
	S. graph	T. time	S. graph size	T. time	T. time	S. graph	T. time	S. graph		
a	41.50	42.85	18406	38.13	1.62	1.84	1.84	2.65	0.051	3680
b	29.62	30.60	15397	28.08	1.31	7.50	7.50	8.60	0.048	7234
c	33.13	34.85	16538	16.87	1.25	0.92	0.92	1.65	0.050	2630
d	26.93	27.61	13548	24.96	1.21	6.55	6.55	6.64	0.041	6973
e	43.16	45.27	19341	37.57	1.24	8.25	8.25	9.28	0.042	8393
f	30.05	31.64	15673	41.78	1.23	7.84	7.84	9.11	0.043	8310
g	33.01	37.40	17234	49.28	1.21	7.20	7.20	7.31	0.042	7589
h	39.03	42.42	18301	43.58	1.27	7.69	7.69	7.73	0.042	7654
i	26.93	27.75	13610	48.15	1.27	5.32	5.32	5.38	0.043	4701
j	14.10	15.19	9461	16.56	1.19	4.61	4.61	5.50	0.045	4768
k	22.65	23.19	12865	41.8	1.32	2.94	2.94	3.80	0.045	3584
l	18.22	21.26	12268	29.86	1.39	2.20	2.20	3.15	0.041	2951
m	6.67	7.07	7186	14.28	1.37	1.34	1.34	1.38	0.042	2601
n	31.63	33.10	16005	30.30	1.25	3.32	3.32	4.19	0.041	3470
o	22.42	23.09	12838	45.98	1.25	17.11	17.11	18.60	0.048	11345
p	18.58	19.17	11258	34.31	1.22	1.25	1.25	1.30	0.050	2544
q	15.73	16.10	10496	30.87	1.30	2.67	2.67	2.71	0.041	3590
r	17.05	21.49	12605	41.44	1.38	2.50	2.50	3.40	0.056	3162
s	42.02	44.26	19289	76.44	1.19	17.37	17.37	18.88	0.048	11470
t	19.73	20.38	11531	27.21	1.22	8.96	8.96	9.96	0.051	8471

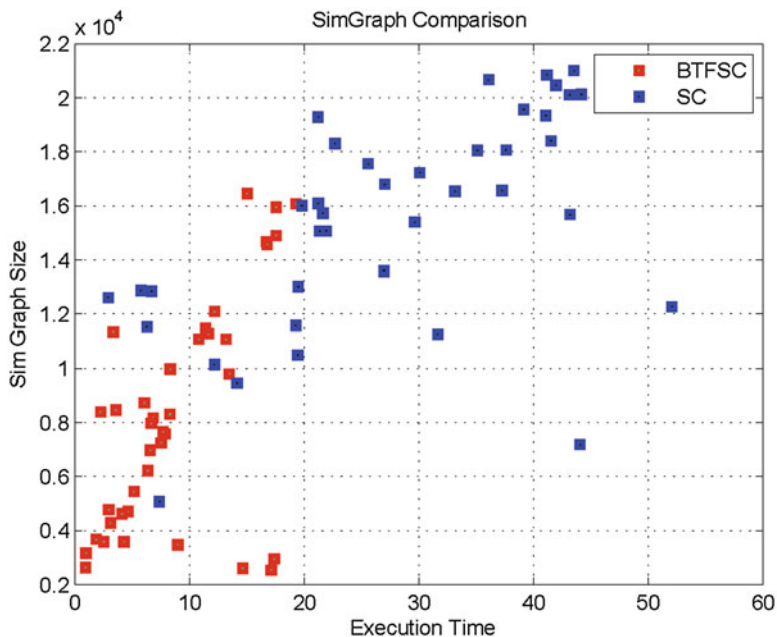


Fig. 7 Execution time vs. similarity graph size. A comparison between proposed method of WBTFSC and spectral clustering for the set of image samples

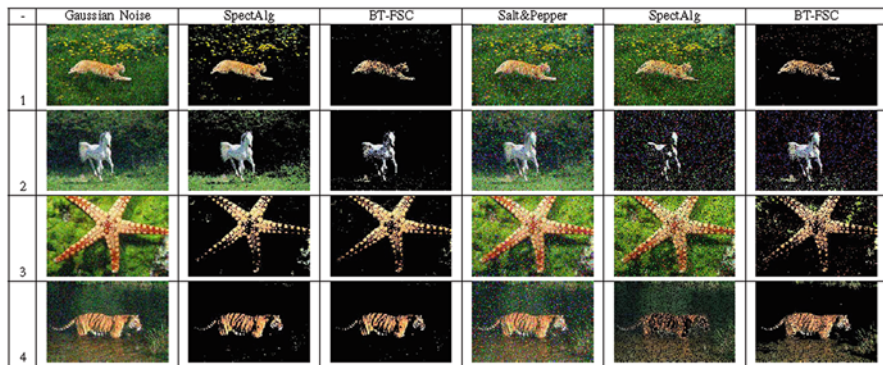
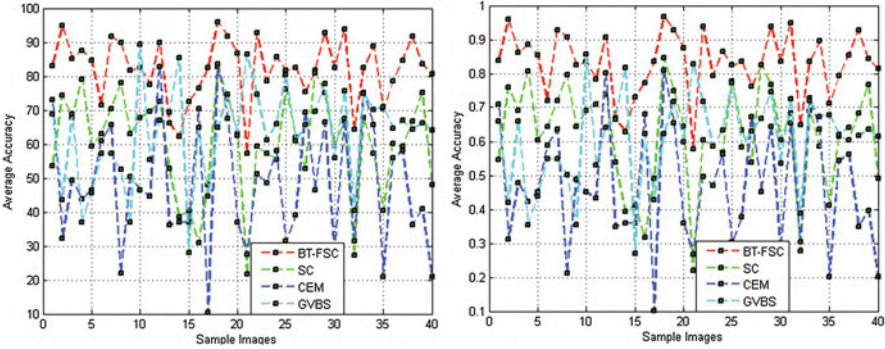


Fig. 8 Visual comparison of noisy images. Gaussian and salt-and-pepper noise are introduced on test images [column 1, column 4], and results can be visually analyzed using spectral algorithm and WBTFSC; (a) original image with Gaussian noise; (b) spectral algorithm; (c) proposed algorithm; (d) original image with salt-and-pepper noise; (e) spectral algorithm; (f) proposed algorithm

where

$$p(R_i, O_{GT}) = \max \left\{ \frac{|R_i \cap O_{GT}|}{|R_i|}, \frac{|R_i \cap O_{GT}|}{|O_{GT}|} \right\} \quad (15)$$



**Fig. 9** Performance measure of proposed approach and three state-of-the-art methods with two different techniques. (Left) proposed approach for calculating accuracy, (right) using approach [25]

For the second approach, consider same test image  $I$  with background  $O'$  where  $O' \in I$  and foreground salient object  $O$  for ( $O \in I$ ). The idea is to compare background of ground truth  $O'_{GT}$ , with the background of proposed algorithm  $O'_{PA}$ , and foreground of ground truth  $O_{GT}$  with foreground of proposed algorithm  $O_{PA}$ . The reason to compare background is to determine the background fragments eradication percentage.

Any change of pixel value greater than selected threshold is considered in calculating efficiency. Percentage is calculated by counting the color variations to the total number of background pixels  $O'_{GT}$  and  $O'_{PA}$ . Foreground efficiency of  $O_{GT}$  and  $O_{PA}$  is calculated by considering all the parameters including the size, shape, and color of the object in ground truth and proposed algorithm. The efficiency results are shown in Fig. 9 (left).

Multiple techniques are adopted to ensure performance measure, e.g., precision and recall (Fig. 10) and mean absolute error (MAE) computed with optimum threshold based on ground truth mask.

Precision and recall is calculated as:

$$\text{Precision} = \frac{\sum_{i=1}^W \sum_{j=1}^H S_{TH}(i, j) S_{GT}(i, j)}{\sum_{i=1}^W \sum_{j=1}^H S_{TH}(i, j)} \quad (16)$$

Recall is calculated as:

$$\text{Recall} = \frac{\sum_{i=1}^W \sum_{j=1}^H S_{TH}(i, j) S_{GT}(i, j)}{\sum_{i=1}^W \sum_{j=1}^H S_{GT}(i, j)} \quad (17)$$



**Table 2** Time parameters for test images with noise. In each partition, “left” column shows WBTFS similarity graph times and total time, whereas “right” column shows spectral similarity graph time and total time

Fig#	Gaussian noise				Salt-and-pepper noise				Speckle noise			
	SG.T	T.T	SG.T	T.T	SG.T	T.T	SG.T	T.T	SG.T	T.T	SG.T	T.T
1	3.27	4.26	52.63	56.56	2.87	3.80	44.68	46.90	2.44	3.81	51.06	53.40
2	18.77	20.59	53.49	57.90	19.18	21.27	41.97	44.21	25.24	27.66	52.29	55.60
3	12.45	14.24	54.64	60.08	8.79	10.164	31.96	33.32	9.28	11.06	49.73	54.13
4	0.40	1.29	55.48	59.95	0.34	1.21	36.28	38.53	0.31	1.33	49.75	54.23

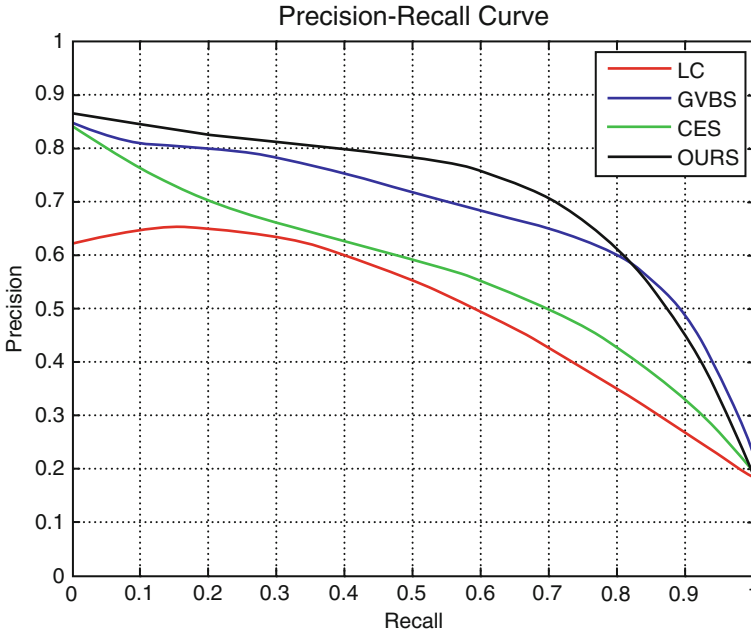


Fig. 10 Precision and recall curve of proposed method on selected datasets

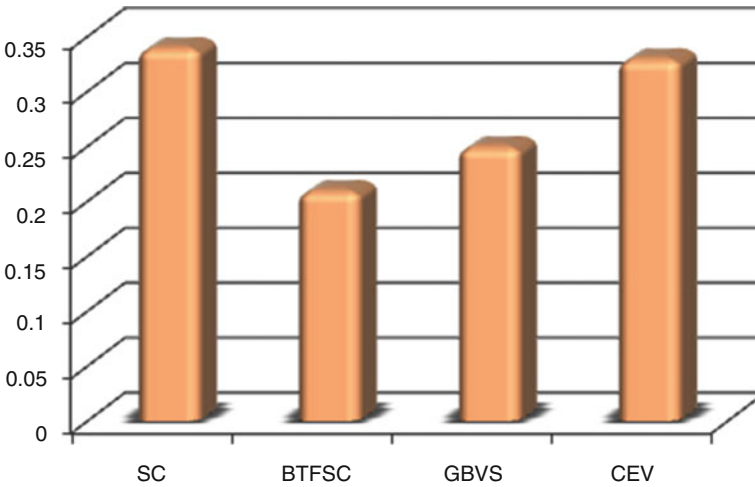


Fig. 11 Mean absolute error of various extraction methods w.r.t. ground truth

where  $S_{TH}$  threshold generated binary object mask corresponds to saliency map and  $S_{GT}$  is its corresponding ground truth (Fig. 11).

For our experiments to calculate precision, we selected fixed thresholding, respectively, for generating binary object mask.

MAE [36] is calculated as:

$$\text{MAE} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H |S_{\text{Extract}}(i, j) - S_{Gt}(i, j)| \quad (18)$$

$W$  and  $H$  are the parameters showing image width and height, respectively. Lower MAE value indicates better performance, which provides a better estimate of dissimilarity between the proposed extraction method and ground truth.

For the proposed method, simulation is done using MATLAB (Ver. 8.0) using Pentium Core 2 Quad processor, 3GHz.

## 6 Chapter Summary

In this chapter, we presented a cascaded design for object of interest detection and finally extraction using weighted binary tree quantization and spectral graph method. The first level of hierarchy reduced distinct number of pixels available in natural scenes using weighted binary tree quantization. The next stage identifies unique color pixels in the image which further reduce number of pixels, as spectral clustering constructs affinity matrix of size  $(n \times n)$ , where  $n$  is the total number of pixels in image.

Selection of unique pixels reduces number of nodes for graph method to overcome memory limitations and computational time. The results are subjected to fuzzy C-means for final grouping, followed by a concluding step of mapping back pixels to their original addresses. The algorithm showed satisfactory results compared to other state-of-the-art algorithms. Noisy images are also tested, Gaussian, salt-and-pepper, and speckle, and achieved amazing results compared to other object extraction approaches. In addition to design's effectiveness, there are few shortcomings which need to be addressed in future articles including number of pre-defined clusters for both weighted binary tree and spectral clustering. Additionally, few conditions should be considered in order to identify the objects of interest as it shows different color and contrast compared to the background, different textural details, etc.

## References

1. Al-Turjman, F. (2018). QoS-aware data delivery framework for safety-inspired multimedia in integrated vehicular-IoT. *Elsevier Computer Communications Journal*, 121, 33–43.
2. Al-Turjman, F., & Alturjman, S. (2018). 5G/IoT-enabled UAVs for multimedia delivery in industry-oriented applications. *Multimedia Tools and Applications*, Springer. <https://doi.org/10.1007/s11042-018-6288-7>.
3. Alabady, S., & Al-Turjman, F. (2018). A novel approach for error detection and correction for efficient energy in wireless networks. *Springer Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-018-6282-0>.

4. Hasan, M. Z., & Al-Turjman, F. (2018). Analysis of cross-layer design of quality-of-service forward geographic wireless sensor network routing strategies in green internet of things. *IEEE Access Journal*, 6(1), 20371–20389.
5. Al-Turjman, F., & Alturjman, S. (2018). Confidential smart-sensing framework in the IoT era. *The Springer Journal of Supercomputing*. <https://doi.org/10.1007/s11227-018-2524-1>.
6. TongKe, F. (2013). Smart agriculture based on cloud computing and IOT. *Journal of Convergence Information Technology*, 8(2), 210–216.
7. Bhatt, C., Dey, N., & Ashour, A. S. (Eds.). (2017). *Internet of things and big data technologies for next generation healthcare* (pp. 978–973). Cham: Springer.
8. Han Zou, Yuxun Zhou, Jianfei Yang, Costas J. Spanos, “Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT”, *Energy and Buildings*, V(174), 309–322, 2018.
9. Ng, A. Y., Jordan, M. I., & Weiss, Y. (2002). On spectral clustering: Analysis and algorithm [C]. NIPS.
10. Hartmann, S. L., & Galloway, R. L. (2000). Depth-buffer targeting for spatially accurate 3-D visualization [J], *medical images. IEEE Transactions on Medical Imaging*, 19(10), 1024–1031.
11. Lei, T., & Sewchand, W. (2001). Object detection and recognition via stochastic model-based image segmentation [C]. *IEEE Multidimensional Signal Processing Workshop*, 2001, Sixth, pp. 17–18.
12. Yixin Chen, J. (2002). A region based fuzzy feature matching a roach to content-based image retrieval [J]. *IEEE Transactions on Fuzzy Systems*, 24(9), 1252–1267.
13. Liu, Z., Shen, L., & Zhang, Z. (2011). Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction [J]. *Signal Processing*, 91(2), 290–299.
14. Fowlkes, C., Belongie, S., Chung, F., & Malik, J. (2004). Spectral grouping using the Nystrom method [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26(2), 214–225.
15. Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259.
16. Mohanta, P. P., Mukherjee, D. P., & ST, A. (2002). Agglomerative clustering for image segmentation [C]. *IEEE International Conference on Pattern Recognition*, 1, 664–667.
17. Zhao, P., & Zhang, C. Q. (2011). A new clustering method and its a lication in social networks [J]. *Pattern Recognition Letters*, 32(15), 2109–2118.
18. Hendrickson, B., & Leland, R.(1995). A multilevel algorithm for partitioning graphs [C]. *IEEE Conference on SuperComputing*.
19. Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 79, 57–70.
20. Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888–905.
21. Yan, D., Huang, L., & Jordan, M. (2009). Fast a roximate spectral clustering [R]: Technical report UCB/EECS]-2009-45.
22. Ducournau, A., Bretto, A., Rital, S., & Laget, B. (2012). A reductive a roach to hypergraph clustering: An a lication to image segmentation [J]. *Pattern Recognition*, 45(7), 2788–2803.
23. Tung, F., Wong, A., & Clausi, D. A. (2010). Enabling scalable spectral clustering for image segmentation [J]. *Pattern Recognition*, 23, 4069–4076.
24. Gao, H., Mei, J., & Si, Y. (2013). A curve evolution a roach for unsupervised segmentation of images with low depth of field [J]. *IEEE Transaction in Image Processing*, 22(10), 4086–4095.
25. Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency [J]. *Advances in Neural Information Processing Systems*, 19, 545–552.
26. Tasdemir, K. (2010). Vector quantization based on proximate spectral clustering of large datasets [J]. *Pattern Recognition*, 45(8), 3034–3044.
27. Zhang, K., Tsang I. W., & Kwok, J. T. (2008) Improved Nystrom low-rank a roximation and error analysis [C]. *International Conference on Machine Learning. ICML* (pp. 1232–1239).
28. Goferman, S., Zelnik-Manor, L., & Tal, A. (2010). Context-aware saliency detection [C]. In *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 2376–2383.

29. Cheng, M. M., Zhang, G. X., Mitra, N. J., & Huang, X. (2011). Global contrast based salient region detection [C]. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 20, 2011.
30. Ng, A. Y., Jordan, M. I., & Weiss, Y. (2001) On spectral clustering: An analysis and an algorithm [J]. *Advance in Neural Information Processing Systems*, 849–856.
31. L.Z. Manor, & P. Perona. (2004). Self tuning spectral clustering [J]. *Advance in Neural Information Processing Systems* (pp. 1601–1608).
32. Luxburg, U. (2007). A tutorial on spectral clustering [J]. *Journal of Statistics and Computing*, 17(4), 395–416.
33. Ge, F., Wang, S., & Liu, T. (2007). New benchmark for image segmentation evaluation [J]. *Journal of Electronic Imaging*, 16(3), 033011.
34. Clinton, N., Holt, A., & Gong, P. (2010). Accuracy assessment measures for object-based image segmentation goodness [J]. *Photogrammetric Engineering & Remote Sensing*, 76(3), 289–299.
35. Duan, Q., Akram, T., Duan, P., & Wang, X. (2016). Visual saliency detection using information contents weighting. *Optik-International Journal for Light and Electron Optics, Elsevier*, 127(19), 7418–7430.
36. Khan, Z. U., Akram, T., Naqvi, S. R., Haider, S. A., Kamran, M., & Muhammad, N. (2018). Automatic detection of plant diseases; utilizing an unsupervised cascaded design. (IBCAST), Islamabad, pp. 339–346, 3028.