Future Computing and Informatics Journal

Volume 2
Issue 2 (2020) *2017, Volume 2, Issue 2*

Article 1

2017

Wavelet based transition region extraction for image segmentation

Priyadarsan Parida
GIET University, priyadarsan.vssut@gmail.com

Nilamani Bhoi
Department of Electronics & Telecommunication Engineering, Veer Surendra Sai University of Technology, Burla 768018, Sambalpur, India, nilamanib@gmail.com

Available online at www.sciencedirect.com

ScienceDirect

Future Computing and Informatics Journal 2 (2017) 65–78
http://www.journals.elsevier.com/future-computing-and-informatics-journal/

Wavelet based transition region extraction for image segmentation

Priyadarsan Parida*, Nilamani Bhoi[1]

Department of Electronics & Telecommunication Engineering, Veer Surendra Sai University of Technology, Burla 768018, Sambalpur, India

Received 9 August 2017; revised 23 September 2017; accepted 24 October 2017
Available online 26 November 2017

Abstract

Transition region based approaches are recent hybrid segmentation techniques well known for its simplicity and effectiveness. Here, the segmentation effectiveness depends on robust extraction of transition regions. So, we have proposed a transition region method which initially decomposes the gray image in wavelet domain. Two existing transition region approaches are applied on approximate coefficients to extract transition region feature matrix. Using this feature matrix the corresponding prominent wavelet coefficients of different bands are found. Inverse wavelet transform are then applied on the modified coefficients to get edge image with more than one pixel width. Otsu thresholding is applied on it to get transition regions. Further, morphological operations are applied to extract the object regions. Finally, the objects are extracted using the object regions. The wavelet domain approach extracts robust transition regions resulting in better segmentation. The proposed method is compared with different existing image segmentation methods. Experimental results reveal that the proposed method achieve 0.95 overall segmentation accuracy. It also works well for extraction of single as well as multiple objects from images.

Keywords: Transition region; Otsu thresholding; Wavelet transform; Edge linking

1. Introduction

Image segmentation is an important pre-processing step for all computer vision and image understanding tasks. It has wide range of applications such as biometrics [1], medical image analysis [2], crop disease detection and classification [3] etc. Image segmentation is the process of separating the object (foreground) from background considering certain features of image such as colour, intensity, texture etc. In the past decade, a wide variety of segmentation techniques are available in literature. In recent era a number of hybrid segmentation techniques have been emerged which provide better segmentation results in comparison to that of the traditional methods. These hybrid segmentations are classified as model based segmentation approaches [4,5], machine learning approaches [6], graph-cut methods [7], active contour and level set methods [8–10] and transition region based approaches [11–18]. In model based approaches, the image is characterised by a statistical model and the model parameters are used as features for segmentation [19]. The machine learning process is basically a training process where a network is trained to optimize the weights of the network from training features of images like texture, brightness etc. After training, the network is presented the query image and it performs classification based segmentation from the tuned weights and learned weights [6]. Graph cut based approaches consider the image as a weighted graph with nodes and vertices, where nodes represent pixels or voxels and vertices represent neighbourhood relationship between pixels. A cost function which represents the cut is optimal in the sense that it effectively separates the object from the background [7]. Active contour based approaches deform its shape in the presence of

* Corresponding author. Fax: +06632430204
E-mail addresses: priyadarsan.vssut@gmail.com (P. Parida), nilamanib@gmail.com (N. Bhoi).
[1] Fax: +06632430204.
Peer review under responsibility of the Faculty of Computers and Information, Future University in Egypt

https://doi.org/10.1016/j.fcij.2017.10.005

external and internal forces leading the contour towards object [8]. In edge-based active contour methods image gradient is used to detect object boundaries [9] whereas region-based active contours use the object and background regions to find an energy optimum where the model fits the image to its best [10]. Level set methods usually make use of an edge indicator to pull the zero level set in the direction of the desired object boundaries [8]. Transition region based methods [11–18] uses transition region for segmentation of images. In local entropy (LE) based method [13] the entropy of a neighbourhood is considered to determine the transition region. It has a limitation that in the event of frequent changes in gray level in a local area, it increases the local entropy and the pixels in the neighborhood is identified as the transition region and it belongs to the foreground or the background. To overcome these disadvantages, Li et al. [14] developed a method for local extraction of the transition region based on the gray level difference (LGLD) which takes into account both changes in gray levels and the extent of these changes. However, the parameter selection unit for determining the threshold value is a problem. The modified local entropy method (MLE) [15] was then put in place to improve the extraction of the transition region. This method also suffers from the same problem as in LGLD. These techniques are ineffective when the foreground and background are of varying intensities. Furthermore, these are mainly used for images that contain a single object. A recent transition region based method named robust single-object image segmentation (RIB) proposed by Zuoyong Li et al. [16] is based on salient transition region provides good segmentation results. But, it only applies to images that contain a single object. In order to mitigate the limitation, Parida et al. [17] proposed local variance and morphological operator based method to yield better performance for both single and multiple object segmentation. In this method the images are categorized into four classes based on whether the background and foreground are simple or textured. The performance of the method degrades when the method is applied on images with (i) textured background and (ii) overlapping gray levels between foreground and background. In order to improve the performance, Parida et al. [18] proposed a novel method using 2-D Gabor filter for transition region. This method works well for overlapping gray levels between background and foreground. The method is unable to perform well when the image contains textured background. Also, it has been found that the proposed method [18] cannot outperform Parida et al. [17] when both the background and foreground are simple. Fig. 3 show the transition regions extraction in the wavelet domain. Fig. 4 shows the segmented output by various methods. From Fig. 4 (d) it is evident that the output is better as compared to other methods. So far the transition region based techniques developed do not perform well when both foreground and background are textured. So, we have proposed a method which overcomes the aforementioned disadvantages. Hence, a method can be developed by taking the common transition region by Refs. [17,18]. The method will work for different types of images in wavelet domain but unable to perform well for images with textured background and foreground as methods [17,18] give inferior

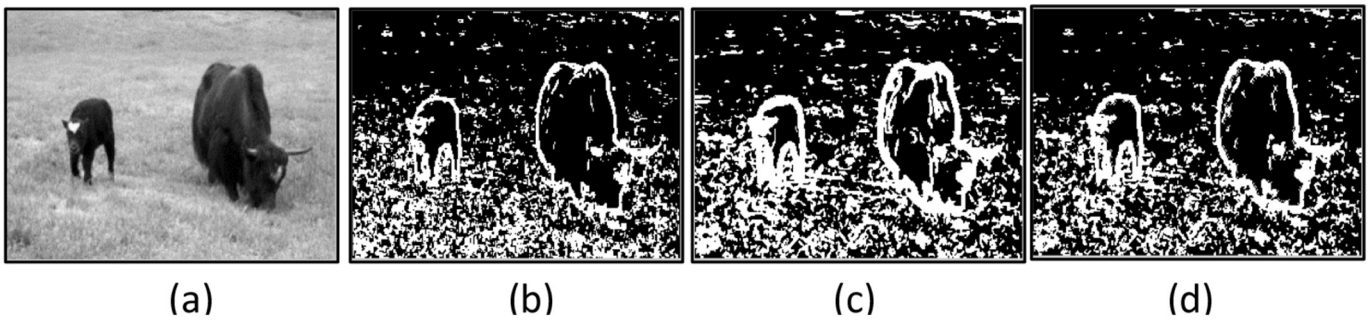

Fig. 1. Transition region extraction of Yak image in spatial domain: (a) Original image, (b) Parida et al. [17], (c) Parida et al. [18], (d) Intersection of (b) and (c) in spatial domain.

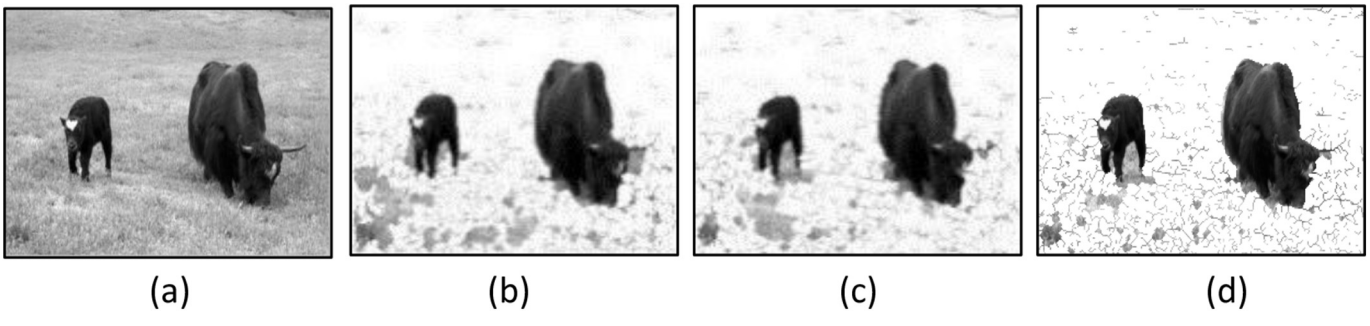

Fig. 2. Segmentation result of Yak image for various methods: (a) Original image, (b) Parida et al. [17], (c) Parida et al. [18], (d) Intersection of (b) and (c) in spatial domain.

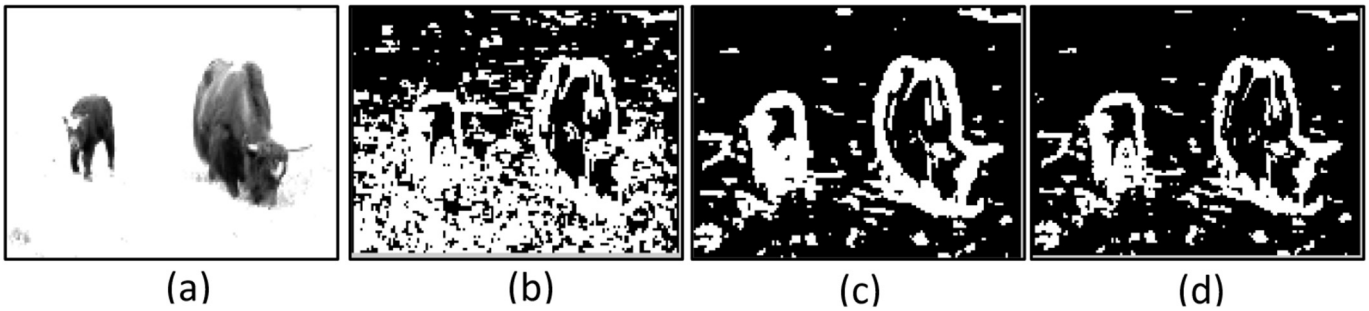

Fig. 3. Transition region extraction of Yak image in wavelet domain: (a) Approximate coefficient image in LL band, (b) Parida et al. [17], (c) Parida et al. [18], (d) Intersection of (b) and (c) in wavelet domain.

Fig. 4. Segmentation result of Yak image for various methods: (a) Original image, (b) Parida et al. [17], (c) Parida et al. [18], (d) Proposed method.

performance when both foreground and background are textured in spatial domain. The transition regions of different methods and the intersection by methods [17,18] are shown in Fig. 1 for Yak image taken from Wisemann dataset [20]. The segmentation results of the corresponding methods are shown in Fig. 2. From Fig. 2 (d) it is seen that backgrounds are still prominent in segmentation result. So far the transition region based techniques developed do not perform well when both foreground and background are textured.

In the proposed method, the image is decomposed in wavelet domain. In the LL band of wavelet domain the foreground coefficients are more prominent as compared to the background. So, the extraction of coefficient corresponding to transition region between background and foreground becomes easier. The methods [17,18] by Parida et al. are applied independently on LL band to find the coefficients. The common approximate coefficients are found using the intersection of the results by the method [17,18]. Using these prominent approx. coefficients of LL band the corresponding coefficients for other bands are retained. Inverse wavelet transform are applied on the modified coefficients to get edge image with more than one pixel width. On the edge image Otsu thresholding is applied to extract transition region. Morphological thinning and edge linking operator are applied to get continuous contour. Morphological region filling operation is applied to get object regions. Finally, the objects are extracted from the object regions.

The rest of the paper is organized as follows: Section 2 describes the existing related work briefly. Section 3 discusses about the proposed method elaborately. Various quantitative measures for detecting the quality of the segmentation are provided in Section 4. Section 5 discusses about the wavelet functions used in the proposed method. The images along with their ground truth used for experimentation is given in Section 6. Section 7 gives result and discussion of proposed method along with other methods. Finally the paper is concluded in Section 8.

2. Related work

The related work required for extraction of transition region for proposed method is discussed below.

2.1. Transition region extraction by Parida et al. [17]

In this approach, the local variance and the mean of the input image is considered for transition region extraction. The local variance is considered as the variance of homogeneous regions differs from regions with edges and fine details. The local variance for a $m \times m$ local neighbourhood is calculated as

$$Lv(i,j) = \sigma^2 = \frac{1}{m^2-1}\sum_{x=1}^{m}\sum_{y=1}^{m}\left(f(x,y)-\overline{f}\right)^2 \quad (1)$$

where, (x,y) denote coordinate in a given neighbourhood of the sub image f and $\overline{f}$ is mean of that neighbourhood. The above process is used throughout the image by sliding the window from left to right and top to bottom to achieve the variance matrix as

$$Lv = \begin{bmatrix} Lv(1,1) & Lv(1,2) & \dots & Lv(1,N) \\ Lv(2,1) & Lv(2,2) & \dots & Lv(2,N) \\ \dots & \dots & \dots & \dots \\ Lv(M,1) & Lv(M,2) & \dots & Lv(M,N) \end{bmatrix} \quad (2)$$

where, M and N correspond to height and width of image. The local variance is compared to a threshold in order to find the transition region according to the equation given below:

$$TR_{Lv}(i,j) = \begin{cases} 1 & if\ Lv(i,j) \geq T_{Lv}; \\ 0 & otherwise, \end{cases} \quad (3)$$

where, TR_{Lv} is the transition region extracted using local variance and T_{Lv} is the threshold which is calculated as the gray level mean of the gray image and it can be determined as

$$T_{Lv} = \frac{1}{M \times N} \sum_{k=1}^{M} \sum_{l=1}^{N} f(k,l) \quad (4)$$

where, f is the original image to be segmented of size $M \times N$.

2.2. Transition region extraction by Parida et al. [18]

In this case, the transition region is extracted by employing a 2-D Gabor filter and statistical standard deviation of Gabor feature image. The image is convolved with a 2-D Gabor filter of 8 orientation to extract the Gabor feature image given as

$$G_f(x,y) = f(x,y) * G(x,y) \quad (5)$$

where, $f(x,y)$ is the gray image and $G(x,y)$ is the impulse response of 2-D Gabor filter. The sign * represent convolution operation. The Gabor kernel $G(x,y)$ is defined by $\psi_{\mu,\nu}(z)$ as

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2} \right) \left[\exp(ik_{\mu,\nu} z) - \exp(-\sigma^2/2) \right] \quad (6)$$

where, the parameters μ and v determine orientation and scale of the Gabor filter kernel. The sign $\|.\|$ denotes norm operator and $z = (x,y)$. The wave vector is denoted by $k_{\mu,\nu} = k_\nu e^{j\varphi_\mu}$ where $k_\nu = k_{\max}/\lambda^\nu$ and $\varphi_\mu = \pi\mu/8$. The spacing between filters in frequency domain is represented by parameter λ. The first term in square brackets determines the oscillation of the kernel, while the second term compensate for the DC term to avoid unwanted depending on the intensity of the filter response of the image. The parameter $k_{\max}$ is set to $\pi/2$ and $\lambda = \sqrt{2}$ [18]. The image is convolved with Gabor filter to generate the Gabor feature matrix as

$$G' = \begin{bmatrix} r(x_0,y_0) & r(x_1,y_0) & . & . & r(x_M,y_0) \\ r(x_0,y_1) & r(x_1,y_1) & . & . & r(x_M,y_1) \\ . & . & . & . & . \\ . & . & . & . & . \\ r(x_0,y_N) & r(x_1,y_N) & . & . & r(x_M,y_N) \end{bmatrix} \quad (7)$$

where, the height and width of the image is given by M and N respectively. Further the feature matrix is normalized using L2-norm as

$$G(x,y) = \|G'(x,y)\| \quad (8)$$

Gabor feature image enhances the image boundary regions of the image. Later thresholding operation is performed to extract the transition regions. In this case the standard deviation of Gabor feature image is considered as threshold. The threshold is determined as

$$T_G = \left(\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (G(x,y) - E)^2 \right)^{\frac{1}{2}} \quad (9)$$

where, E is the mean value given as

$$E = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} G(x,y) \quad (10)$$

The transition region is extracted from Gabor feature image using T_G as

$$TR_G(x,y) = \begin{cases} 1 & if \quad G(x,y) > T_G \\ 0 & if \quad G(x,y) \leq T_G \end{cases} \quad (11)$$

where, TR_G is the transition region extracted using 2-D Gabor filter and thresholding.

3. Proposed method

The proposed method starts with decomposition of the original gray image in wavelet domain. In LL band the foreground is more prominent supressing the background textures. This leads to easier extraction of transition region between foreground and background. The methods [17,18] are applied on LL band and intersected for identifying the prominent features Further, employing the abovementioned technique the detailed prominent coefficients are retained. Inverse wavelet transform is applied on the modified approx. and detailed coefficients to get an edge image with more than one pixel width. Further, Otsu thresholding is employed on the edge image to extract transition regions. Morphological thinning, edge-linking and region filling operation is employed on the transition region to extract the object regions. Finally, objects are extracted via these object regions. Fig. 5 shows the architecture of the proposed method.

3.1. Wavelet decomposition of image

Wavelets transform is an effective tool for time-frequency analysis. Various wavelet family exist in literature [21,22]. The input image is subjected to a two dimensional discrete wavelet transform (DWT). The image is decomposed into approximate, horizontal, vertical and detailed coefficient matrices. The approximate coefficients represent low frequency components whereas the detailed coefficients (horizontal, vertical and diagonal) represent the high frequency

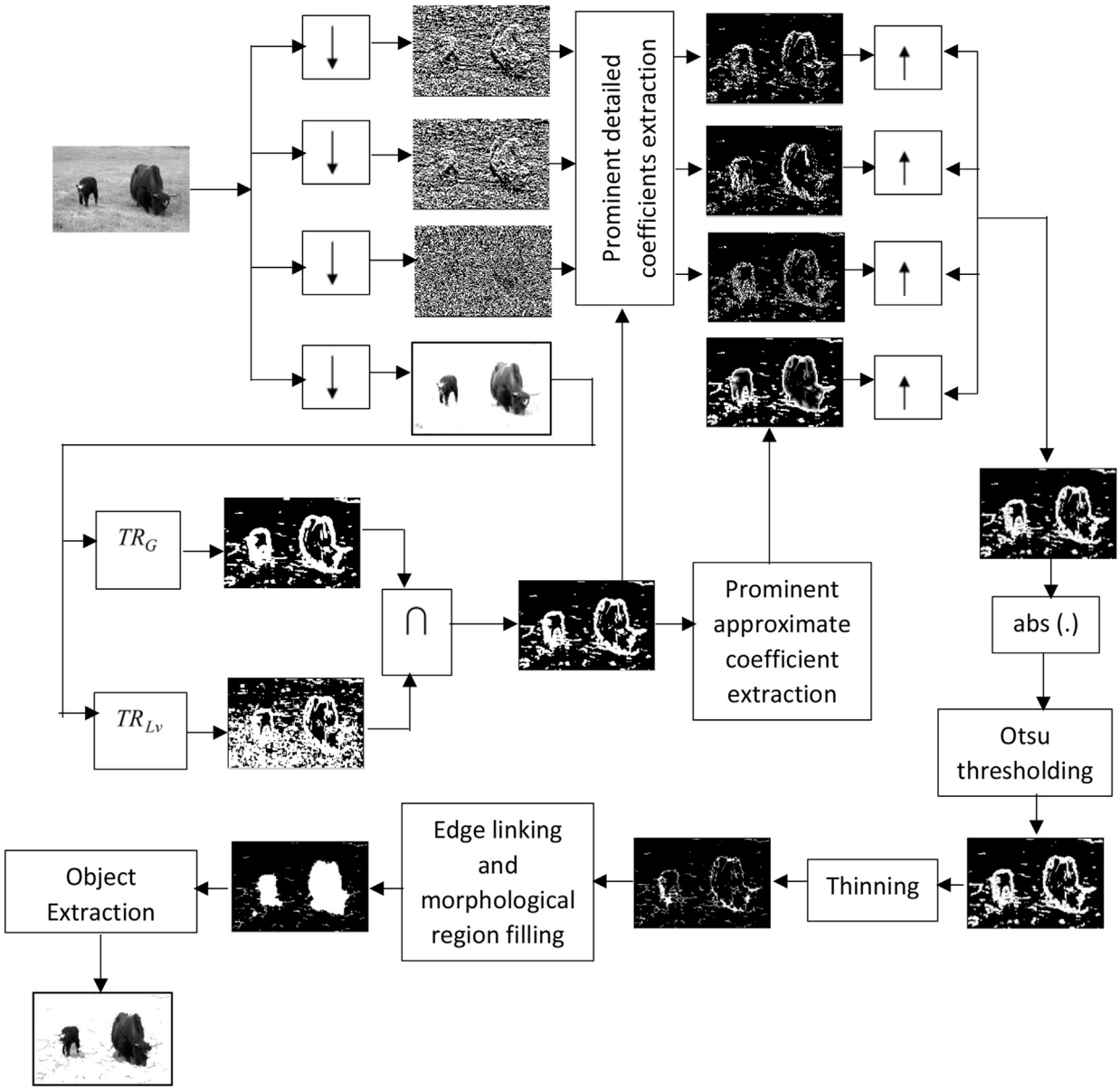

Fig. 5. Architecture of the proposed method for image segmentation.

components which provides both high frequency values for both background and foreground (i.e. both edge and none edge regions). This can be better visualized from Fig. 6.

3.2. Extraction of high frequency information of object regions

In decomposition process from the former step, the high and low frequency components and the background as well as the foreground texture information are discriminated. The next goal is to retain prominent coefficients corresponding to the object regions. So, to extract these coefficients near the object boundary, the LL band is subjected to object boundary extraction using two transition methods [17,18]. The former considers local variance of the approximate coefficients given by TR_{Lv} whereas the later takes care of the texture portion or uneven illumination arises due to the smoothing process given by TR_G. The calculation of TR_{Lv} and TR_G is done in the same manner as that of the methods discussed in Section 2. These methods results in two coefficient images of the approximate coefficient matrix. The intersection of the results determine the exact boundary region between the object and background.

The intersection matrix consists of 1 and 0, where 1 represent the pixel surrounding the object. Using the intersection values the new approximate and detail (horizontal, vertical and diagonal) coefficient matrices are extracted. The new detail coefficients represent only high frequency components of the boundary regions surrounding the objects.

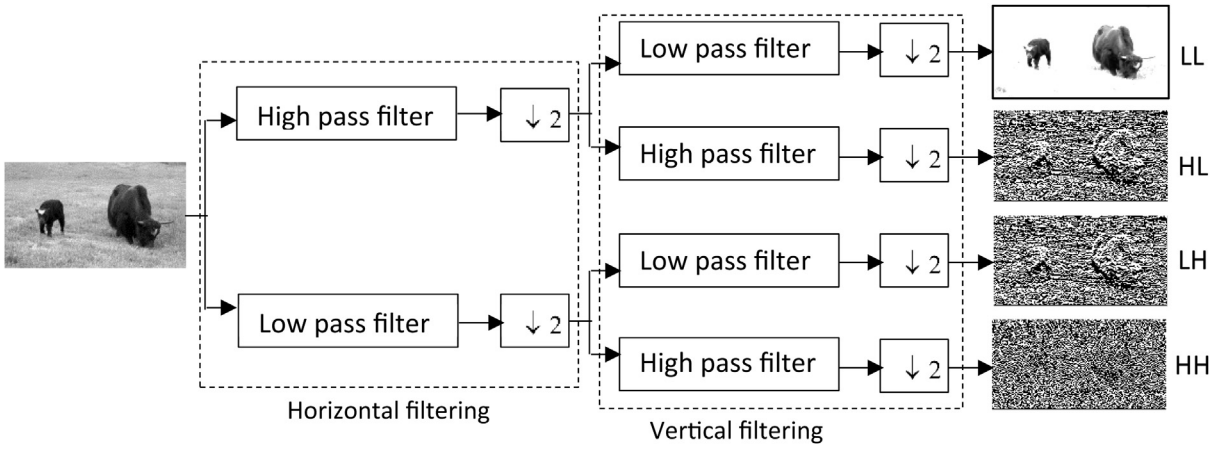

Fig. 6. Decomposition of Yak image using two dimensional discrete wavelet transform (DWT).

3.3. Reconstruction of wavelet feature image and extraction of transition region

The wavelet feature image is reconstructed from the new detailed coefficients as well as the new approximate matrix obtained from the former step using inverse discrete wavelet transform. The same wavelet function used for decomposition is used for reconstruction. Basically we have used the 1-level decomposition and reconstruction using Haar wavelet function. The selection of particular Haar wavelet for decomposition as well as reconstruction is discussed briefly in Section 5. The reconstructed feature image may have some negative coefficients so, we consider its magnitude part to extract the reconstructed wavelet feature image. Finally, from the feature image we extract the transition region by applying the Otsu threshold on feature image.

3.4. Morphological thinning and edge linking operation

The transition regions of more than one pixel width are generated from the former step. These transition region undergo morphological thinning operation to yield object contour of single pixel width. But during thinning operation, it is possible that the edge image achieved do not have all edges of the object connected or continuous. Therefore, the edge image further undergoes edge linking operation. Various edge linking approaches exists in literature [23,24]. The binding of edge is applied to the pixel when the pixel has 8-connectivity or 8-adjacent connection or pixel is not associated with any other edges.

The edge linking process is carried out as [16]: On each endpoint (intermittent edge pixel), 8-connectivity is checked to the end point, or junction is encountered. A number is assigned as a label for all endpoints along with the junctions. The distance between adjacent endpoints or junctions are calculated. If the city block distance between adjacent endpoints or junctions are less than 10, then the points are linked [17]. The unlabeled isolated points are left out in the process. Finally, in the edge linking process a continuous edge of the object region is achieved which is further called as object contour.

3.5. Morphological region filling operation

The object contour obtained in the former step is a closed in nature. It is further operated with morphological region filling operation. The morphological region filling operation results in the extraction of object regions from the object contours. The structuring element used for morphological operation is of disk type with radius 3.

3.6. Object extraction

The object region extracted from the previous step is a binary image where the object regions are marked with value 1 and background as 0. This also referred as segmentation mask. From the segmentation mask the segmented object is extracted by replacing the 1 with the original gray value from the input image and the background 0 is replaced by a value 255.

4. Performance measures

The performance of the proposed method is quantitatively evaluated via 5 different performance measures: misclassification error (ME) [25,26], false positive rate (FPR) [26], false negative rate (FNR) [26], Jaccard index (JI) [27] and segmentation accuracy (SA) [28].

In a binary classification, the foreground (object) pixels erroneously classified as background or vice versa is termed as misclassification error (ME) [25]. The ME is defined as

$$ME = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|} \tag{24}$$

where, B_O and F_O denote the background and object pixels of ground truth image whereas B_T and F_T represent the background and foreground pixels of the segmentation result. The operator $|.|$ represent cardinality of set operation. The ME

value lie between 0 and 1 where 0 represent complete segmentation without any deviation from ground truth. The value 1 correspond to totally erroneous segmentation result. The less the value of ME correspond to better segmentation result. The measures FPR and FNR define the former measure more precisely.

The FPR is the number of background pixels classified as object pixels to the total number of background pixel. The FNR is number of pixels in the object classified in the background pixels to total object pixels. The FPR and FNR is defined as

$$FPR = \frac{|B_O \cap F_T|}{|B_O|} \tag{25}$$

$$FNR = \frac{|F_O \cap B_T|}{|F_O|} \tag{26}$$

The value of FPR and FNR also varies between 0 and 1. The lower the value of FPR and FNR, the better is segmentation result. Higher values of FPR and FNR makes the result to be highly over segmentation and under segmentation respectively.

To evaluate the similarity of the segmentation result with the ground truth Jaccard index is used. The Jaccard index [27] is defined as

$$JI = \frac{|GT \cap SR|}{|GT \cup SR|} \tag{27}$$

where, GT and SR correspond to ground truth and segmentation result respectively. The JI value varies between 0 and 1. Higher value (i.e., close to 1) denote better segmentation result or maximum resemblance with the ground truth (required segmentation result).

Segmentation accuracy (SA) [28] is a global measure which denote the ratio of total well classified pixels in the segmentation result which is given as

$$\text{SA} = \frac{\text{Number of correctly segmented pixels}}{\text{Total number of pixels}} \tag{28}$$

The value of SA remains in the range from 0 to 1. High SA value indicate better segmentation accuracy. Based on the above five performance measures the proposed method is quantitatively compared with various segmentation methods.

5. Choice of wavelet for effective segmentation

The various parameters for extracting TR_{Lv}, TR_G are taken as per [17,18] respectively. The choice of proper wavelet basis function depends upon various applications such as enhancement, de-noising, texture feature extraction etc. The proposed method is tested by applying Haar, Daubechies, Coiflets, Symlet and Bi-orthogonal wavelet families with different decomposition levels were tested against the performance measures (ME, FPR, FNR, JI and SA) for different images. For single level decomposition Haar, Daubechies, Coiflets and Symlet basis functions outperforms well. This can be well verified from Table 1 where we have determined the various performance measures of the proposed method for Kid and Dog image. For both Kid and Dog image the proposed method achieves best ME, FPR, FNR using single level decomposition for Haar, Daubechies, Coiflets and Symlet wavelet family. As Haar wavelet is simplest we have considered this in our proposed method for effective segmentation.

6. Images and ground truths

For experimentation, we have considered 12 images from Wisemann [20] and MSRM [29] dataset. The images considered can be classified into 4 broad categories depending on the type of foreground/background to be simple or textured. The different categorization of images are given in Table 2. Fig. 7 represents the original images with their corresponding ground truths.

7. Result and discussion

The entire experiment is performed in a PC with Core-i3, 1.9 GHz processor and 8 GB RAM. The experimentation is done in MATLAB 7.0 environment. The proposed method is compared with several existing transition region based methods such as LE [13], MLE [15], RIB [16], Parida et al. [17] and Parida et al. [18] along with a recent level set based approach to image segmentation with image inhomogeneity (LSISIM) [30]. The images used for experimentation are 8-bit and of different resolution to demonstrate the effectiveness of proposed method. The images contain both single and multiple objects in the experimentation process.

To begin our analysis, the images along with their corresponding ground truths considered for experimentation are taken from Wisemann dataset [20] and MSRM database [29]. The images contains both single and multiple images with both foreground/background simple as well as textured. For each image we have computed five quantitative measures such as ME, FPR, FNR, JI and SA for various methods along with our proposed method which are listed in Table 3. The best values for different methods for every image appear as bold. The qualitative evaluation can be verified from the segmentation results and the segmentation masks shown in Figs. 8 and 9 respectively.

For Boat image, the proposed method attains lowest value of ME, FNR and highest JI, SA respectively. The method [17] attains lowest FPR indicating that the result of the proposed method is a little over-segmented than [17] which can be visualized from Fig. 9. Similarly for Dog image, the proposed method outperforms except for FNR. The method RIB achieves the lowest FNR for Dog image representing the result of proposed method to be bit under-segmented. The Kid image is having uniform texture distribution throughout so its best performance measures are provided by proposed method except for its FNR. The lowest FNR for Kid image is provided by Ref. [17]. For Mushroom image, the proposed method provide the best performance measures except for FNR. The method RIB has the lowest FNR. From Fig. 9 the result of

Table 1
Performance measures of Dog and Kid images for different wavelet basis functions and levels.

Sl.No.	Image	Wavelet	Number of tap	ME	FPR	FNR	JI	SA
1	Dog	Daubachies	1	0.0288	0.0201	0.0689	0.8527	0.9712
		Daubachies	2	0.0511	0.0393	0.1047	0.7589	0.9489
		Daubachies	3	0.0647	0.0508	0.1280	0.7075	0.9353
		Daubachies	4	0.0840	0.0708	0.1444	0.6465	0.9160
		Haar	1	0.0288	0.0201	0.0689	0.8527	0.9712
		Bi-orthogonal	1.1	0.0288	0.0201	0.0689	0.8527	0.9712
		Bi-orthogonal	1.3	0.0666	0.0622	0.0870	0.7109	0.9334
		Bi-orthogonal	1.5	0.0763	0.0743	0.0854	0.6826	0.9237
		Bi-orthogonal	2.2	0.0529	0.0446	0.0906	0.7553	0.8606
		Bi-orthogonal	2.4	0.0663	0.0610	0.0902	0.7113	0.9337
		Coiflets	1	0.0621	0.0476	0.1283	0.7161	0.9379
		Coiflets	2	0.1006	0.0899	0.1493	0.6030	0.8994
		Symlets	1	0.0288	0.0201	0.0689	0.8527	0.9712
		Symlets	2	0.0511	0.0393	0.1047	0.7589	0.9489
		Symlets	3	0.0647	0.0508	0.1280	0.7075	0.9353
		Symlets	4	0.0830	0.0713	0.1367	0.6512	0.9170
2	Kid	Daubachies	1	0.0342	0.0216	0.0931	0.8233	0.9658
		Daubachies	2	0.0456	0.0127	0.1993	0.7560	0.8611
		Daubachies	3	0.0547	0.0171	0.2302	0.7130	0.9453
		Daubachies	4	0.0795	0.0513	0.2111	0.6367	0.9205
		Haar	1	0.0342	0.0216	0.0931	0.8233	0.9658
		Bi-orthogonal	1.1	0.0342	0.0216	0.0931	0.8233	0.9658
		Bi-orthogonal	1.3	0.0506	0.0291	0.1508	0.7479	0.9494
		Bi-orthogonal	1.5	0.0519	0.0288	0.1594	0.7410	0.9481
		Bi-orthogonal	2.2	0.0488	0.0162	0.2009	0.7431	0.9512
		Bi-orthogonal	2.4	0.0521	0.0237	0.1843	0.7345	0.9479
		Coiflets	1	0.0508	0.0123	0.2301	0.7281	0.9492
		Coiflets	2	0.0649	0.0373	0.1933	0.6872	0.9351
		Symlets	1	0.0342	0.0216	0.0931	0.8233	0.9658
		Symlets	2	0.0456	0.0127	0.1993	0.7560	0.9544
		Symlets	3	0.0547	0.0171	0.2302	0.7130	0.9453
		Symlets	4	0.0698	0.0326	0.2432	0.6571	0.9302

Table 2
Classification of various images considering the complexity of foreground and background.

Category	Foreground	Background	Images
Category-1	Simple	Simple	Player Papaya
Category-2	Simple	Textured	Rock Signboard Dustbin
Category-3	Textured	Simple	Dog Kid Boat
Category-4	Textured	Textured	Yak Rock Puppy Mushroom

Mushroom image for RIB shows that although it contains two separate objects it shows a single connected object region (both the objects connected by background) which indicates that the method falsely detects the background varying texture as object region. The Papaya image is a simple image with portion of foreground textured. For it the proposed method achieves lowest ME, JI and SA whereas the method RIB and [17] attains lowest FPR and FNR respectively. For Player image, the proposed method provides the lowest value of ME, FPR and highest values of JI, SA. The method LSISIM provides the lowest FNR indicating that the method suffers from low under-segmentation than the proposed method. Visually, it can be verified from Fig. 9 that, the segmentation mask represents much higher background region in comparison to that of the proposed method which is not desirable. The Rock and Yak images are having a uniform texture variation in their background. The best values of ME, SA and JI is obtained by the proposed method for both Yak and Rock image indicating the result to be more similar and accurate in comparison to the ground truths. The method RIB attains a lowest FPR for Rock image. From Fig. 8, visually it can be implied that the method RIB loses majority foreground region and background regions due to which it provides a low FPR value. Similarly, the lowest value of FNR is provided by Ref. [17], indicating that the proposed method loses some object information to achieve more background suppression in comparison to [17]. The similar case can be observed from Table 3 for Signboard image except for the FNR value. The Player image is a complex image where the proposed method achieves the best ME, JI and SA. For player image the lowest FPR and FNR is provided by Ref. [17] and LE respectively. Visual description from Figs. 8 and 9 describe that segmentation result for LE provides more background information along with object

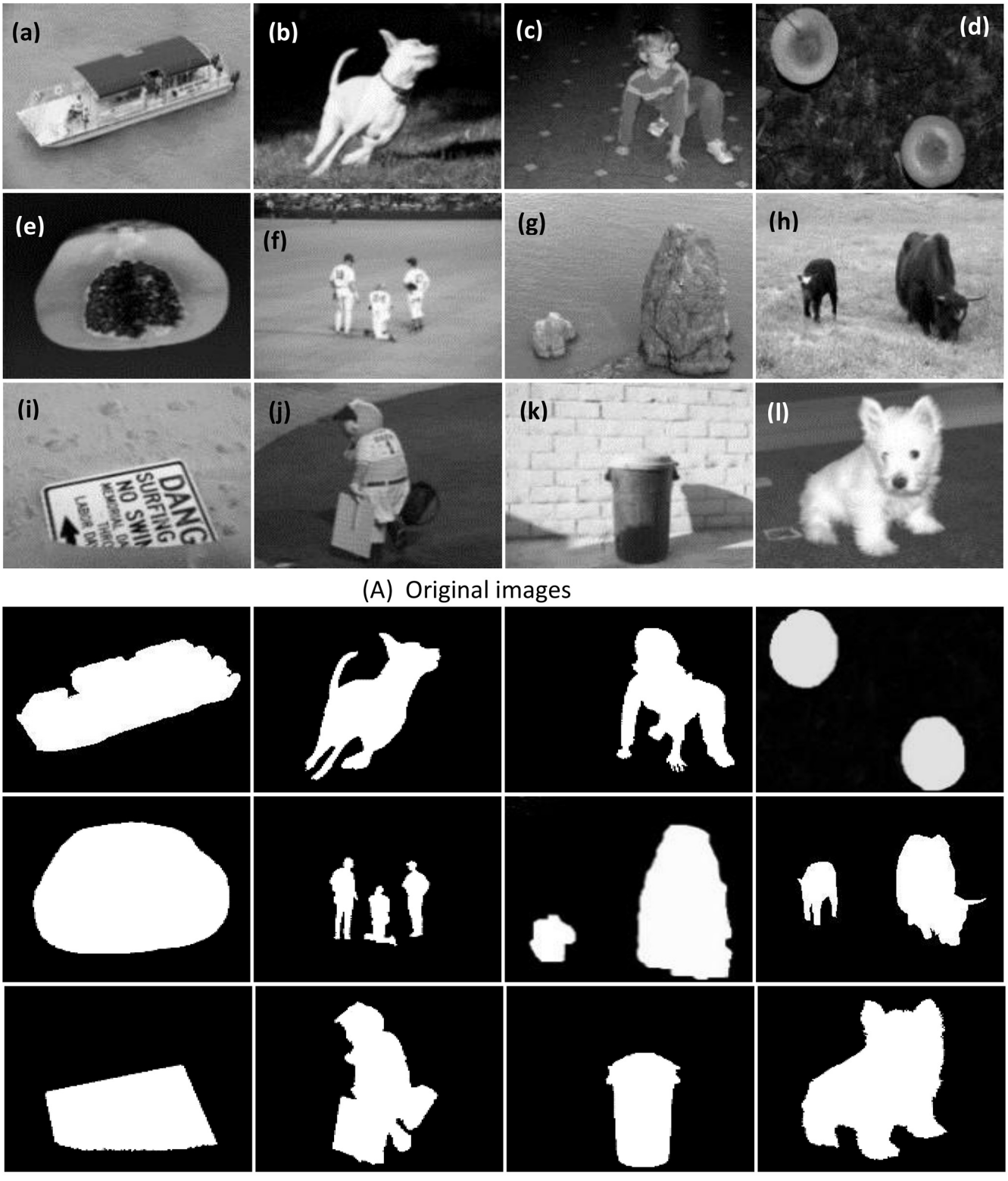

Fig. 7. (A) Original images: (a) Boat, (b) Dog, (c) Kid, (d) Mushroom, (e) Papaya, (f) Players, (g) Rock, (h) Yak, (i) Signboard, (j) Player, (k) Dustbin, (l) Puppy; (B) Ground truths.

Table 3
Performance measures (ME, FPR, FNR, JI, SA) of different methods for various types of images.

Sl.no.	Image	Methods	ME	FPR	FNR	JI	SA
1	Boat	LE	0.2435	0.2658	0.1916	0.4993	0.7565
		MLE	0.1871	0.0436	0.5243	0.4316	0.8129
		RIB	0.1784	0.0121	0.5688	0.4193	0.8216
		LSISIM	0.2267	0.0897	0.5459	0.3756	0.7733
		Parida et al. [17]	0.0622	**0.0017**	0.2023	0.7936	0.9378
		Parida et al. [18]	0.0578	0.0021	0.1875	0.8086	0.9422
		Proposed method	**0.0388**	0.0040	**0.1197**	**0.8723**	**0.9612**
2	Dog	LE	0.8387	0.9496	0.3319	0.1251	0.1613
		MLE	0.6678	0.6480	0.7588	0.0607	0.3322
		RIB	0.2131	0.2571	**0.0113**	0.4535	0.7869
		LSISIM	0.1168	0.1220	0.0930	0.5822	0.8832
		Parida et al. [17]	0.2476	0.2958	0.0273	0.4135	0.7524
		Parida et al. [18]	0.0901	0.1004	0.0432	0.6559	0.9099
		Proposed method	**0.0288**	**0.0201**	0.0689	**0.8527**	**0.9712**
3	Kid	LE	0.6351	0.7354	0.1678	0.1879	0.3649
		MLE	0.3121	0.2609	0.5524	0.2012	0.6879
		RIB	0.1171	0.0364	0.4959	0.4306	0.8829
		LSISIM	0.1597	0.0902	0.4838	0.3634	0.8403
		Parida et al. [17]	0.1363	0.1541	**0.0531**	0.5510	0.8637
		Parida et al. [18]	0.0437	0.0352	0.0833	0.7873	0.9563
		Proposed method	**0.0342**	**0.0216**	0.0931	**0.8233**	**0.9658**
4	Mushroom	LE	0.7866	0.9265	0.1114	0.1624	0.2134
		MLE	0.8191	0.8116	0.8549	0.0296	0.1809
		RIB	0.7974	0.9634	**0.0001**	0.1777	0.2026
		LSISIM	0.2157	0.2557	0.0227	0.4375	0.7843
		Parida et al. [17]	0.6353	0.7648	0.0101	0.2110	0.3647
		Parida et al. [18]	0.1911	0.2249	0.0281	0.4661	0.8089
		Proposed method	**0.0717**	**0.0773**	0.0447	**0.6969**	**0.9283**
5	Papaya	LE	0.5925	0.7658	0.3835	0.3205	0.4075
		MLE	0.3266	0.0638	0.6455	0.3291	0.6734
		RIB	0.3204	**0.0000**	0.7093	0.2907	0.6796
		LSISIM	0.0372	0.0226	0.0548	0.9202	0.9628
		Parida et al. [17]	0.0199	0.0153	**0.0254**	0.9570	0.9801
		Parida et al. [18]	0.0195	0.0122	0.0284	0.9576	0.9805
		Proposed method	**0.0157**	0.0056	0.0279	**0.9655**	**0.9843**
6	Players	LE	0.3179	0.3310	0.1492	0.1612	0.6821
		MLE	0.1710	0.1439	0.5243	0.1656	0.8290
		RIB	0.1557	0.0909	1.0000	0.0000	0.8443
		LSISIM	0.5146	0.5526	**0.0242**	0.1199	0.4854
		Parida et al. [17]	0.0824	0.0807	0.1048	0.4382	0.9176
		Parida et al. [18]	0.0979	0.0950	0.1357	0.3880	0.9021
		Proposed method	**0.0754**	**0.0702**	0.1426	**0.4493**	**0.9246**
7	Rock	LE	0.4803	0.6678	**0.0014**	0.3690	0.5197
		MLE	0.2069	0.0989	0.4820	0.4138	0.7931
		RIB	0.1867	0.0242	0.6003	0.3765	0.8133
		LSISIM	0.2157	0.2557	0.0227	0.4375	0.7843
		Parida et al. [17]	0.1809	0.0507	0.5136	0.4307	0.8191
		Parida et al. [18]	0.0890	0.0834	0.1034	0.7391	0.9110
		Proposed method	**0.0801**	**0.0165**	0.2436	**0.7258**	**0.9199**
8	Yak	LE	0.8063	0.9283	0.1559	0.1419	0.1937
		MLE	0.2388	0.1918	0.4904	0.2517	0.7612
		RIB	0.1312	**0.0102**	0.7781	0.2104	0.8688
		LSISIM	0.0820	0.0441	0.0304	0.5849	0.9080
		Parida et al. [17]	0.1373	0.1597	**0.0182**	0.5303	0.8627
		Parida et al. [18]	0.1234	0.1381	0.0452	0.5500	0.8766
		Proposed method	**0.0497**	0.0482	0.0581	**0.7481**	**0.9503**
9	Signboard	LE	0.4402	0.5333	0.1432	0.3172	0.5598
		MLE	0.2046	0.1256	0.4572	0.3872	0.7954
		RIB	0.0245	0.0290	0.0100	0.9060	0.9755
		LSISIM	0.3287	0.3620	0.2222	0.3609	0.6713
		Parida et al. [17]	0.0458	0.0591	**0.0033**	0.8385	0.9542
		Parida et al. [18]	0.0225	0.0204	0.0293	0.9114	0.9775
		Proposed method	**0.0103**	**0.0040**	0.0304	**0.9575**	**0.9897**
10	Player	LE	0.1752	0.2011	**0.0751**	0.5205	0.8248
		MLE	0.1814	0.0455	0.7093	0.2470	0.8186
		RIB	0.1911	0.0144	0.8778	0.1157	0.8089
		LSISIM	0.8400	0.8918	0.6401	0.0810	0.1600
		Parida et al. [17]	0.0405	0.0120	0.1509	0.8116	0.9595
		Parida et al. [18]	0.0405	0.0215	0.1142	0.8179	0.9595
		Proposed method	**0.0353**	**0.0113**	0.1283	**0.8352**	**0.9647**
11	Dustbin	LE	0.6327	0.6993	0.2205	0.1463	0.3673
		MLE	0.2597	0.1676	0.8348	0.0807	0.7403
		RIB	0.1642	**0.0348**	0.9722	0.0228	0.8358
		LSISIM	0.1435	0.1496	**0.1055**	0.4644	0.8565
		Parida et al. [17]	0.2080	0.0890	0.9447	0.0357	0.7920
		Parida et al. [18]	0.2019	0.0868	0.9145	0.0556	0.7981
		Proposed method	**0.0799**	0.0729	0.1235	**0.6030**	**0.9201**
12	Puppy	LE	0.3686	0.4111	0.2777	0.3846	0.6314
		MLE	0.3379	0.0787	0.8941	0.0906	0.6621
		RIB	0.3169	**0.0070**	0.9821	0.0177	0.6831
		LSISIM	0.1083	0.1351	0.0512	0.7363	0.8917
		Parida et al. [17]	0.3074	0.0201	0.9211	0.0756	0.6926
		Parida et al. [18]	0.0354	0.0443	**0.0162**	0.8987	0.9646
		Proposed method	**0.0213**	0.0213	0.0211	**0.9361**	**0.9787**

region producing a low under-segmented result in comparison to that of proposed method. The Dustbin image is a highly textured image consisting of shadow of the object in the background and the contrast variation of background is highly non-uniform. The proposed method achieves the lowest ME value whereas highest JI and SA value indicating that the result is much more similar to the ground truth. The method RIB and LSISIM provide lowest FPR and FNR respectively for Dustbin image. It can be clearly visualized from Fig. 8 it that RIB misses majority object region along with background producing low FPR value. Similarly, LSISIM adds up more background along with object providing low FNR value. The puppy image is simple image having a slow contrast variation near the object edge regions. The proposed method attains lowest ME whereas highest JI and SA. The RIB and [18] provide lowest FPR, FNR respectively. Visual comparison from Fig. 8 reveal that RIB results contains object along with more background pixels making it low under segmented values than the proposed method.

To show the effectiveness of the proposed method in comparison with other methods their average performance measure is calculated for all images and are given in Table 4. The best values of the performance measures appear as bold. From Table 4 it is evident that the proposed method has the lowest average ME, FPR and FNR whereas it has highest JI and SA. Hence, the proposed method can be treated as better method as compared to others.

The proposed method is tested against different variety of images from BSDS dataset [31] where we have considered the images with overlapping objects in different environments. The segmentation results are given in Fig. 10. In case of overlapping multi object images, when the objects are overlapping in nature, it is found that the proposed method considers the multiple object case as a single object and segment the multiple object unlike a single object segmentation.

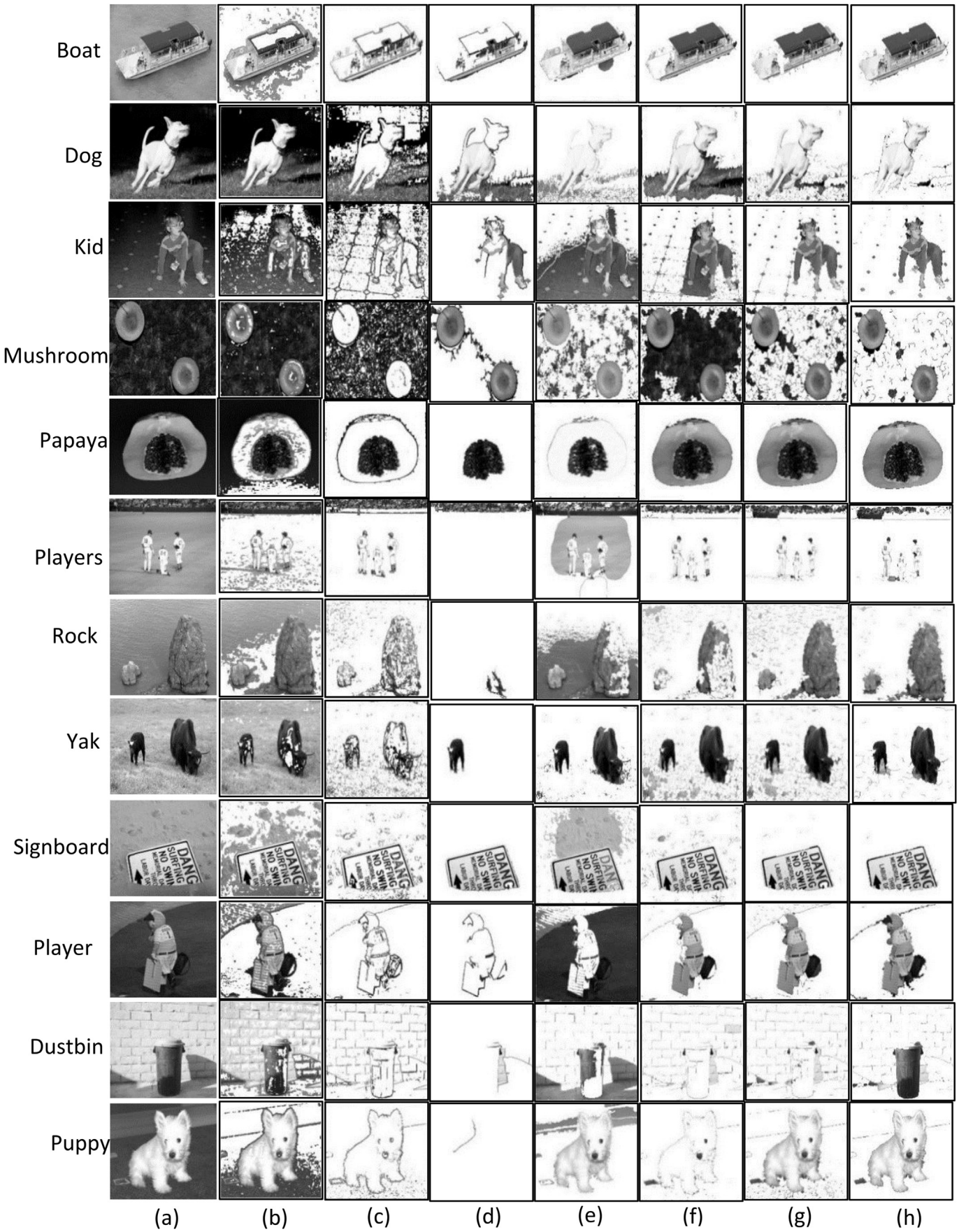

Fig. 8. Segmentation result of various methods for different images (a) Original gray image, (b) LE, (c) MLE, (d) RIB, (e) LSACM, (f) Parida et al. [17], (g) Parida et al. [18], (h) Proposed method.

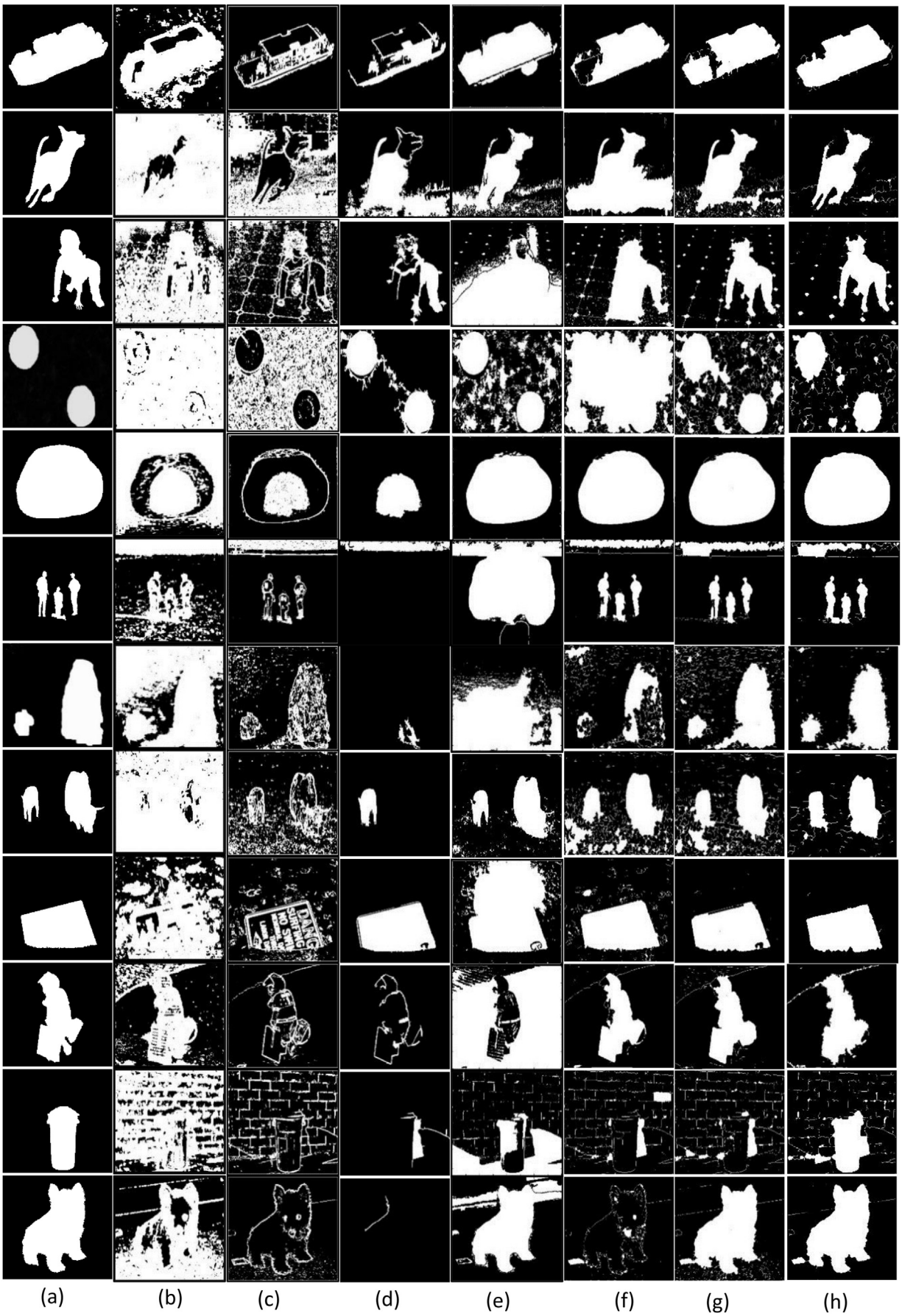

Fig. 9. Segmentation masks (a) Ground truth, (b) LE, (c) MLE, (d) RIB, (e) LSACM, (f) Parida et al. [17], (g) Parida et al. [18], (h) Proposed method.

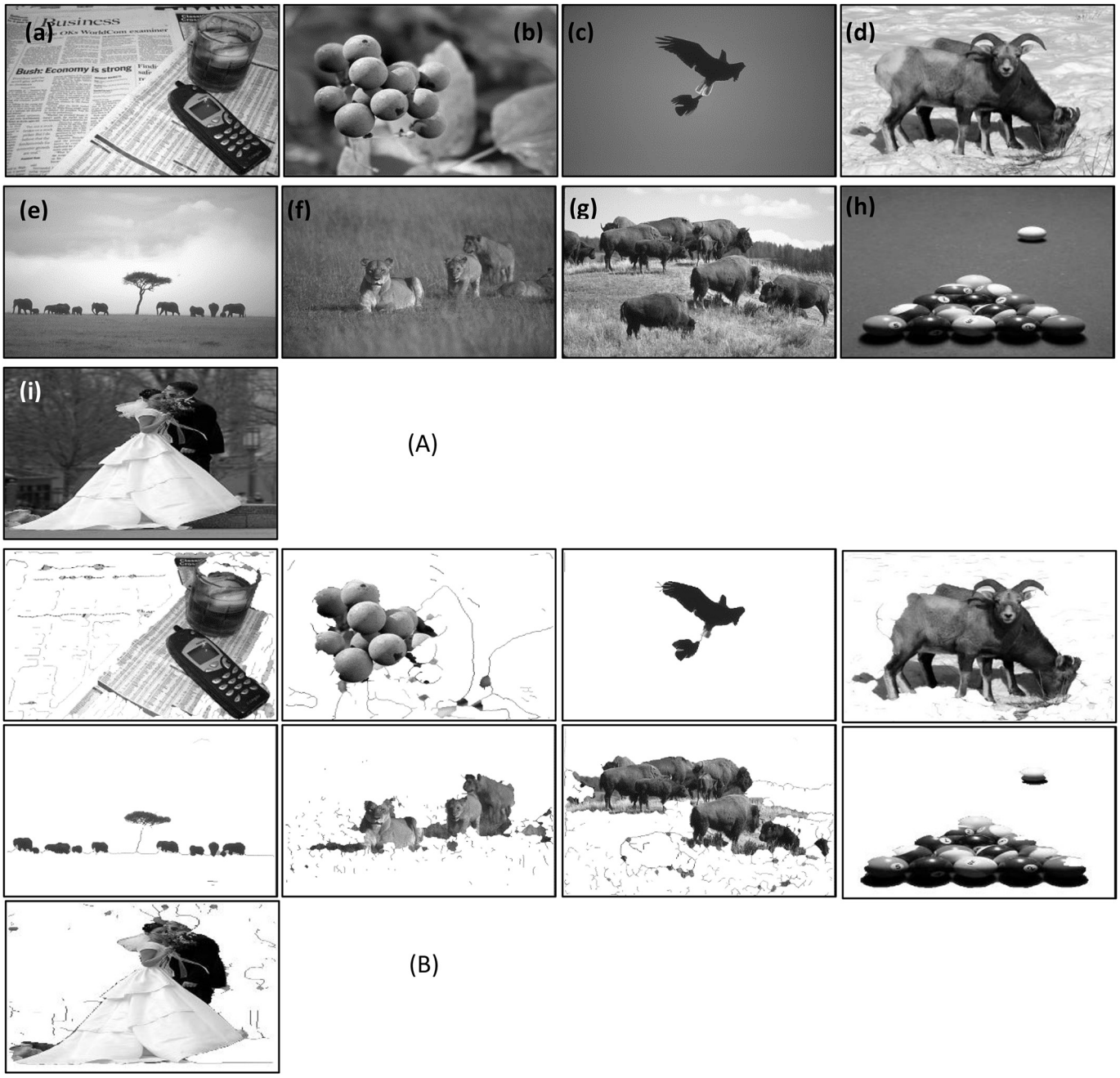

Fig. 10. Segmentation results of images containing multiple overlapping objects: (A) Original images: (a) Cellphone, (b) Berries, (c) Bird, (d) Goats, (e) Elephants, (f) Lions, (g) Grazing cattle, (h) Billiard board, (i) Couple; (B) Segmented result of each image using the proposed method.

Table 4
Average performance of different methods for various performance measures.

Method	Average ME	Average FPR	Average FNR	Average JI	Average SA
LE	0.5265	0.6179	0.1841	0.2780	0.4735
MLE	0.3261	0.2233	0.6440	0.2241	0.6739
RIB	0.2331	0.1233	0.5838	0.2851	0.7669
LSISIM	0.2491	0.2476	0.1914	0.4553	0.7501
Parida et al. [17]	0.1753	0.1419	0.2479	0.5072	0.8247
Parida et al. [18]	0.0844	0.0720	0.1441	0.6697	0.9156
Proposed method	**0.0451**	**0.0311**	**0.0918**	**0.7888**	**0.9549**

8. Conclusion

The paper presents a new transition region extraction method using wavelet transform for image segmentation. The transition region extracted is robust in the sense that the proposed method is able to extract the exact transition region correctly for all types of images (i.e. for both foreground and background simple as well as textured). Approximate coefficients of the image decomposed in wavelet domain contain supressed background and foreground texture features. When the texture features are supressed, the coefficients of

background/foreground regions become prominent as compared to foreground/background regions. This helps in better extraction of transition regions which in turn leads to better segmentation. So, the developed method can extract the single as well as multiple objects from the images which can be verified from the experimental results. The obtained result has also less emergence of background and low loss of object regions.

References

[1] Mondal S, Bours P. Neurocomputing A study on continuous authentication using a combination of keystroke and mouse biometrics. Neurocomputing 2017;230:1–22. https://doi.org/10.1016/j.neucom.2016.11.031.

[2] Iii WMW, Mondal S, Bours P, Akram T, Naqvi SR, Haider SA, et al. Medical Image Analysis – past, present, and future. Comput Electr Eng 2017;59:1–22. https://doi.org/10.1016/j.neucom.2016.11.031.

[3] Akram T, Naqvi SR, Haider SA, Kamran M. Towards real-time crops surveillance for disease classification : exploiting parallelism in computer vision R. Comput Electr Eng 2017;59:15–26. https://doi.org/10.1016/j.compeleceng.2017.02.020.

[4] Xia Y, Ji Z, Zhang Y. Brain MRI image segmentation based on learning local variational Gaussian mixture models. Neurocomputing 2016;204:189–97. https://doi.org/10.1016/j.neucom.2015.08.125.

[5] Ji Z, Xia Y, Sun Q, Chen Q, Feng D. Adaptive scale fuzzy local Gaussian mixture model for brain MR image segmentation. Neurocomputing 2014;134:60–9. https://doi.org/10.1016/j.neucom.2012.12.067.

[6] Ren Malik. Learning a classification model for segmentation. Proc Ninth IEEE Int Conf Compu Vision IEEE 2003;1:10–7. https://doi.org/10.1109/ICCV.2003.1238308.

[7] Boykov YY, Jolly M-P. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Proceedings eighth IEEE international conference on computer vision. ICCV 2001, vol. 1. IEEE comput. Soc; 2001. p. 105–12. https://doi.org/10.1109/ICCV.2001.937505.

[8] Zhou Y, Shi W-R, Chen W, Chen Y, Li Y, Tan L-W, et al. Active contours driven by localizing region and edge-based intensity fitting energy with application to segmentation of the left ventricle in cardiac CT images. Neurocomputing 2015;156:199–210. https://doi.org/10.1016/j.neucom.2014.12.061.

[9] Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. Int J Comput Vis 1988;1:321–31. https://doi.org/10.1007/BF00133570.

[10] Chen Y, Zhang J, Mishra A, Yang J. Image segmentation and bias correction via an improved level set method. Neurocomputing 2011;74:3520–30. https://doi.org/10.1016/j.neucom.2011.06.006.

[11] Gerbrands JJ. Segmentation of noisy images. Delft (Netherlands): Technische Univ.; 1988.

[12] Zhang YJJ, Gerbrands JJJ. Transition region determination based thresholding. Pattern Recognit Lett 1991;12:13–23. https://doi.org/10.1016/0167-8655(91)90023-F.

[13] Yan C, Sang N, Zhang T. Local entropy-based transition region extraction and thresholding. Pattern Recognit Lett 2003;24:2935–41. https://doi.org/10.1016/S0167-8655(03)00154-5.

[14] Li Z, Liu C. Gray level difference-based transition region extraction and thresholding. Comput Electr Eng 2009;35:696–704. https://doi.org/10.1016/j.compeleceng.2009.02.001.

[15] Li Z, Zhang D, Xu Y, Liu C. Modified local entropy-based transition region extraction and thresholding. Appl Soft Comput J 2011;11:5630–8. https://doi.org/10.1016/j.asoc.2011.04.001.

[16] Li Z, Liu G, Zhang D, Xu Y. Robust single-object image segmentation based on salient transition region. Pattern Recognit 2016;52:317–31. https://doi.org/10.1016/j.patcog.2015.10.009.

[17] Parida P, Bhoi N. Transition region based single and multiple object segmentation of gray scale images. Eng Sci Technol Int J 2016;19:1206–15. https://doi.org/10.1016/j.jestch.2015.12.009.

[18] Parida P, Bhoi N. 2-D Gabor filter based transition region extraction and morphological operation for image segmentation. Comput Electr Eng 2016;0:1–16. https://doi.org/10.1016/j.compeleceng.2016.10.019.

[19] Kim SC, Kang TJ. Texture classification and segmentation using wavelet packet frame and Gaussian mixture model. Pattern Recognit 2007;40:1207–21. https://doi.org/10.1016/j.patcog.2006.09.012.

[20] Alpert S, Galun M, Brandt A, Basri R. Image segmentation by probabilistic bottom-up aggregation and cue integration. IEEE Trans Pattern Anal Mach Intell 2012;34:315–27. https://doi.org/10.1109/TPAMI.2011.130.

[21] Bakhshipour A, Jafari A, Nassiri SM, Zare D. Weed segmentation using texture features extracted from wavelet sub-images. Biosyst Eng 2017;157:1–12. https://doi.org/10.1016/j.biosystemseng.2017.02.002.

[22] Khan MM, Mendes A, Zhang P, Chalup SK. Evolving multi-dimensional wavelet neural networks for classification using Cartesian Genetic Programming. Neurocomputing 2017;247:39–58. https://doi.org/10.1016/j.neucom.2017.03.048.

[23] Chen Q, Sun Q sen, Ann Heng P, Xia D shen. A double-threshold image binarization method based on edge detector. Pattern Recognit 2008;41:1254–67. https://doi.org/10.1016/j.patcog.2007.09.007.

[24] Zhijie W, Hong Z. Edge linking using geodesic distance and neighborhood information. In: IEEE/ASME international conference on advanced intelligent mechatronics, AIM; 2008. p. 151–5. https://doi.org/10.1109/AIM.2008.4601650.

[25] Yasnoff WA, Mui JK, Bacus JW. Error measures for scene segmentation. Pattern Recognit 1977;9:217–31. https://doi.org/10.1016/0031-3203(77)90006-1.

[26] Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. J Electron Imaging 2004;13:146. https://doi.org/10.1117/1.1631315.

[27] Tizhoosh HR, Othman AA. Anatomy-aware measurement of segmentation accuracy. In: Styner MA, Angelini ED, editors. Proc. SPIE 9784, medical imaging 2016: image processing; 2016. https://doi.org/10.1117/12.2214869. 97840C.

[28] Rajaby E, Ahadi SM, Aghaeinia H. Robust color image segmentation using fuzzy c-means with weighted hue and intensity. Digit Signal Process 2016;51:170–83. https://doi.org/10.1016/j.dsp.2016.01.010.

[29] Liu T, Yuan Z, Sun J, Wang J, Zheng N, Tang X, et al. Learning to detect a salient object. IEEE Trans Pattern Anal Mach Intell 2011;33:353–67. https://doi.org/10.1109/TPAMI.2010.70.

[30] Zhang K, Zhang L, Lam K-M, Zhang D, Member S, Lam K-M, et al. A level set approach to image segmentation with intensity inhomogeneity. IEEE Trans Cybern 2016;46:546–57. https://doi.org/10.1109/TCYB.2015.2409119.

[31] Martin D, Fowlkes C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Proc Eighth IEEE Int Conf Comput Vis ICCV 2001 2001;2:1–15. https://doi.org/10.1109/ICCV.2001.937655.