

Detection of Multiple Salient Objects by Categorizing Regional Features

Kang-Han Oh, Soo-Hyung Kim*, Young-Chul Kim and Yu-Ra Lee

School of Electronics and Computer Engineering Chonnam National University, 77 Yongbong-ro, Buk-gu,
Gwangju, 500-757, Korea

[e-mail: blastps@naver.com, shkim@jnu.ac.kr, yckim@jnu.ac.kr, ebfk3751@naver.com]

*Corresponding author: Soo-Hyung Kim

*Received May 22, 2015; revised August 10, 2015; accepted October 31, 2015;
published January 31, 2016*

Abstract

Recently, various and effective contrast based salient object detection models to focus on a single target have been proposed. However, there is a lack of research on detection of multiple objects, and also it is a more challenging task than single target process. In the multiple target problem, we are confronted by new difficulties caused by distinct difference between properties of objects. The characteristic of existing models depending on the global maximum distribution of data point would become a drawback for detection of multiple objects. In this paper, by analyzing limitations of the existing methods, we have devised three main processes to detect multiple salient objects. In the first stage, regional features are extracted from over-segmented regions. In the second stage, the regional features are categorized into homogeneous cluster using the mean-shift algorithm with the kernel function having various sizes. In the final stage, we compute saliency scores of the categorized regions using only spatial features without the contrast features, and then all scores are integrated for the final salient regions. In the experimental results, the scheme achieved superior detection accuracy for the SED2 and MSRA-ASD benchmarks with both a higher precision and better recall than state-of-the-art approaches. Especially, given multiple objects having different properties, our model significantly outperforms all existing models.

Keywords: Multiple salient objects detection, Object segmentation, Clustering algorithm

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(NRF-2015R1D1A3A01018993) and this research was financially supported by the Ministry of Education, Science Technology (MEST) and National Research Foundation of Korea (NRF) through the Human Resource Training Project for Regional Innovation (NRF-2014H1C1A1066771).

1. Introduction

The human visual attention ability is an important process that allows individuals to concentrate on important and attractive objects in complex scenes. The saliency detection aims to implement the human visual attention system. Various saliency detection models have been implemented in a wide range of image processing and pattern recognition applications, including object detection [1-2], visual tracking [3], object recognition [4], image summarization [5], and image/video compression [6].

In the past decade, many previous models have been exploited for detecting the saliency region. Especially, regional feature based approaches [20-21] have achieved excellent performance against others in the single object images, because these approaches utilized the regional contrast features to focus on the most highlight region and pre-segmentation technique for object region preserving (i.e graph based segmentation or super-pixel segmentation) however several limitations are commonly observed in the multiple object images. The limitations are related to distinct properties of objects. An example is shown in Fig. 1.

For the first limitation in the green circles (Fig. 1), prior methods lose object regions when there is a big difference between contrast values of objects. The contrast features have been widely used for detecting salient region [7,8,9,10,20,21,22], because it is a reasonable feature representing the human visual attention system that automatically focuses on uniqueness and rarity. However, the contrast features are not suitable for the multiple object problem because they tend to value the most highlight region. This property can significantly reduce energy on the other objects during the normalization process. The red circles in Fig. 1 describe the second limitation of the existing methods that tend to under-value small objects. Even though simple background images, they are missing a small object because the prior methods have used post-processing to develop their saliency detection results using their pre-detected saliency clues representing an estimated centroid of object. Obviously, this process, which can focus on the global maximum point, is highly effective at the single object problem but this process has an adverse effect on the detection of the multiple objects, which are widely located in the whole image.

To overcome these limitations, the proposed method was designed to include two properties. Firstly, we have adopted only spatial features, which represent the spatial variation and location prior, as an objectness without contrast features while computing the final saliency score. Secondly, in the homogeneous region classification part, we have utilized various kernel sizes while implementing a non-parametric clustering algorithm to counteract objects of various sizes.

This paper is organized as follows. In Section 2, we briefly demonstrate related works with a description of their advantages and disadvantages. The details of the proposed model are described in Section 3. In Section 4, the experimental results of the proposed method are evaluated over the MSRA-ASD and SED2 benchmarks against the state-of-the-art approaches. In Section 5, we present a conclusion and future work.

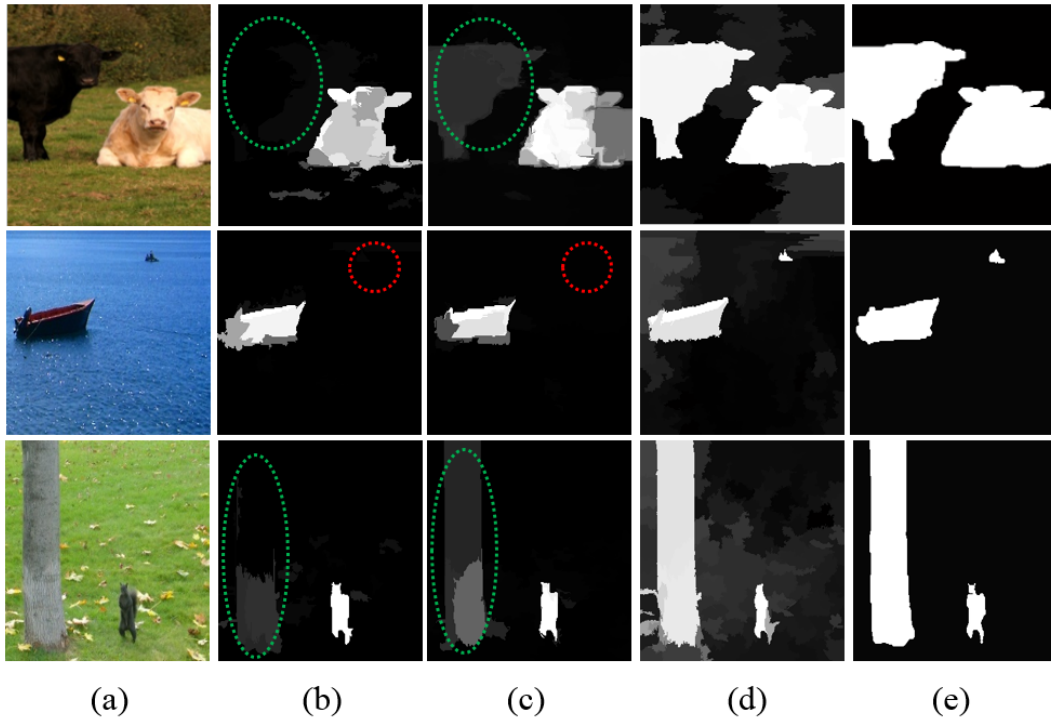


Fig. 1. Limitations of previous models with multiple objects. (a) Original, (b) RC [20], (c) DRFI [21], (d) Our work, (e) Ground truth.

2. Related Work

Over the past few decades, a considerable number of studies for saliency detection have been proposed based on different paradigms. The paradigms can be classified into three categories, such as global contrast, local contrast and regional feature based method.

In general, the main idea of the global contrast based method is to utilize the mean pixel energy value of the entire image to compute the contrast of each pixels on the image. Achanta et al. [7] extracted the maximal symmetric area and proposed novel dissimilarity metrics to compute the saliency. An image contrast symmetry approach was proposed in [8], which used color, orientation and contrast symmetry features for computing the saliency map. The method by Rosin et al. [9] computes the saliency map using edges in a grayscale image, threshold decomposition, and distance transform. Bruce and Tsotsos proposed saliency detection method using an Independent Component Analysis (ICA) containing local image patches from large trained dataset [10]. These models successfully detect the salient region in the simple background images, however bad global contrast and complex background of image can reduce their performance, and the resulting saliency map heavily depend on most highlight region.

The local contrast (or window-based) based models to detect saliency have been introduced by researchers to overcome global contrast method problems by analyzing local neighborhoods of pixels. Rahtu et al. [14] introduced a saliency detection method, which combines saliency measures based on the conditional random field (CRF). The illumination, color and motion information are used for applying a sliding window to the image, and the local features are computed between the distribution of the inner window and the external

window. Then, in order to recover a well-defined salient object, the estimated saliency is used in the CRF model to define an energy minimization-based segmentation approach. Klein et al. [15] presented a new approach for computing saliency by using a Kullback-Leibler divergence. They determined the contrast of the center and the surrounding distribution of features for color, intensity, and orientation. Although these approaches are useful in detecting saliency, the limitation is that the results have the local maximum error in computing saliency, and the methods have to modify the window size for different objects at different scales.

Recently, the regional features based methods [20], [21], adopted the pre-segmentation approaches (graph-based, super-pixel) to detect salient regions. These approaches have the advantages of region preserving and texture use. Chen et al. [20] proposed a global contrast saliency (RC) based on regional values. This method is effective in the simple background images but the limitations are observed in complex background images, which have various patterns. Jiang et al. [21] introduced a saliency detection method (DRFI) by combining regional contrast vectors representing color and texture unlike RC, which uses traditional contrast value. Because of these feature of DRFI we can assume that it can reach outstanding performance in the complex background images, which have various type of textures and colors. However the DRFI is supervised learning-based model, which requires a lot of training time and selecting dataset for training so the performance of the DRFI significantly depend on training dataset.

3. Proposed Method

In this section, we describe the proposed method to compute multiple salient objects. The main idea is that after obtaining over-segmented image, we extract regional features, which are Color, Edge and Spatial, and then the regional features are categorized into homogeneous categories using the mean-shift algorithm with various kernel functions. Finally, given the categorized images, we compute regional saliency scores based on spatial variation and overlapping ratio with image borders without the contrast features. The proposed framework is described in Fig. 2.

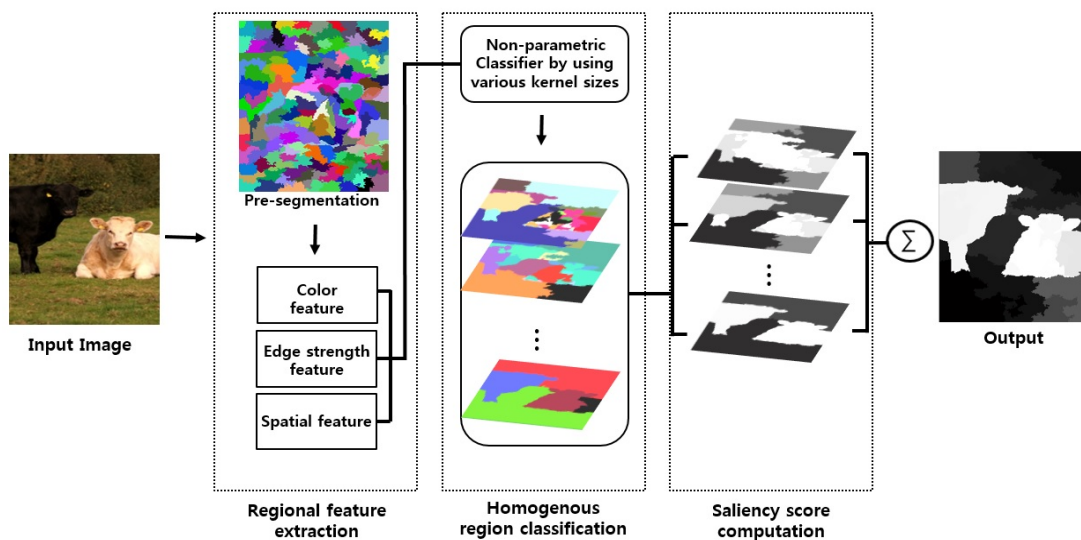


Fig. 2. Framework of proposed method.

3.1 Regional feature extraction

Given an image I , we generate over-segmentation image $I = \{ R_1, R_2, \dots, R_n \}$ using super-pixel algorithm [24]. The fourteen regional features, which represent color, edge and spatial distribution, are extracted from the overlapping regions between each feature map f_Map and observed super-pixel R_i . For this reason, each super-pixel region provides fourteen vectors $R_i = \{ f_1, f_2, \dots, f_{14} \}$. A framework of the regional feature extraction is clearly illustrated in Fig. 3. Here, an equation of the regional feature extraction is given by

$$f = \frac{1}{h_R^d} \sum_{(x,y) \in R} f_Map_{xy} \quad (1)$$

Where R is observed super-pixel, (x, y) is coordinates, f_Map represents the feature map and $\frac{1}{h_R^d}$ means the area of the observed R .

As the color features, we adopt the RGB, HSV, and LAB. The reason for taking three color models is to counteract various conditions of input images. Each color model includes different features. Firstly, the LAB color space designed to approximate human vision is widely used for salient region detection. Secondly, the HSV is invariant to certain types of highlights, shading and shadows. Finally, the RGB has advantages with its simplicity and feature of preserving color information. The advantage of edge strength features is that its results are highly effective when either background or foreground is a simple. For the edge features, we convert gray scale image into edge map using the canny and sobel edge detection algorithms. The converted edge maps are utilized for computing edge strength representing frequency of edge pixels in the observed super-pixel. The edges are important features since they may correspond to significant features of objects in the scene and provides solid local texture patterns to classify vectors into homogeneous categories in brightness changes environment. For the spatial feature, the Gaussian spatial weighting map located on image center and centroid coordinates corresponding to each categorized region are used. While implementing the mean-shift algorithm, the spatial features provide location information of observed vector and prevent that the categorized regions, which have similar properties, are widespread in the spatial domain. Table 1 shows fourteen dimensional features used in this work.

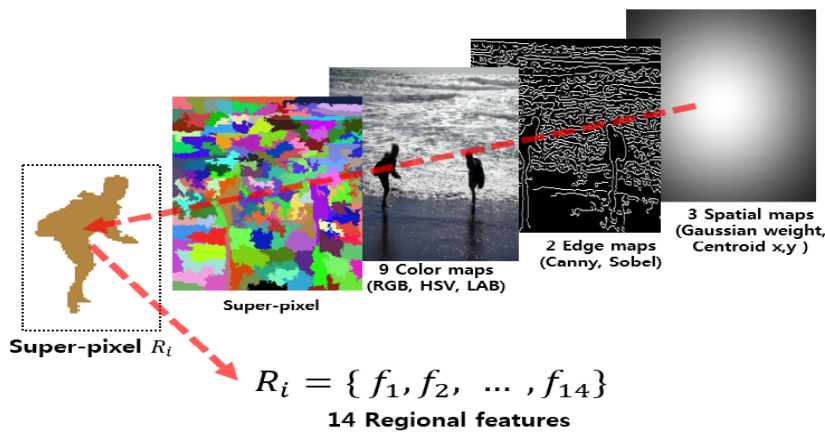


Fig. 3. Proposed regional feature extraction.

Table 1. The features used in this work.

	Features (dimension)	Dimension
Color feature	RGB (3),HSV (3),LAB (3)	9
Edge strength feature	Canny(1), Sobel(1)	2
Spatial feature	Gaussian energy(1), centroid coordinates(2)	3

3.2 Homogeneous region classification

The mean-shift algorithm [17] is a useful non-parametric clustering technique, which does not require prior knowledge about number of clusters and does not constrain the shape of the clusters. However the mean-shift algorithm is compute-intensive, requiring in general $O(kN^2)$ operations, where N is number of input vectors, and k is the number of average integrations for each input vector. In our work, we consider only 150 region vectors by computing super-pixel algorithm, and this strategy significantly reduces time consumption of mean-shift compared to considering whole pixel information on the image. Before computing mean-shift, all feature vectors are normalized. We estimate the joint density with the kernel,

$$K(x) = \left\{ \frac{e}{h_s^{sd} h_r^{rd}} k_E \left(\left\| \frac{x^s}{h_s} \right\|^2 \right) k_E \left(\left\| \frac{x^r}{h_r} \right\|^2 \right) \right\} \quad (2)$$

where rd is dimension of color and edge features and the spatial dimension is sd , where x^s is the spatial components, which are centroid vectors and Gaussian weight map. The x^r represents the color vectors and edge strength. The h_s and h_r are the corresponding kernel bandwidth, and e the corresponding normalization constant so that $\sum_{u=1}^m K_u(x) = 1$, where $k_E(x)$ is the Epanechnikov kernel profile. An advantage of the Epanechnikov kernel is that its estimated density is smooth so the kernel function is quite solid with local noises [17].

$$k_E(x) = \begin{cases} 1 - x & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases} \quad (3)$$

To categorize regional features, we have applied various kernel sizes while implementing mean-shift. The reason for taking differing kernel sizes is that there is no single method to predict optimal condition of various input images for successful clustering. For example, if input image includes big objects, the large scale kernel is optimal condition, whereas the small scale kernel is appropriate for small objects. As can be seen in Fig. 4, small objects are eliminated by the large kernel in the middle image, and the case of the small kernel has detected the small objects but the biggest object is over-segmented in the right image. In the mean-shift process, the detection of the small, and big objects can be considered as to find the local maximum and global maximum points respectively. Therefore, we adopt the various sizes of the kernel profiles to robustly detect the differing object sizes in the scene. While computing mean-shift, the two kernel sizes representing spatial feature and others are gradually increased until there is no variation of saliency scores (or number of clusters ≥ 3) and the Fig. 6 first row describes this process.



Fig. 4. Visualization result by the different kernel scale. From left to right : original image, categorized result with large kernel (size : 0.95), categorized result with small kernel (size : 0.05).

3.3 Saliency score computation

In our work, the main characteristic to define salient region is ultimately determined in this section. Given the vectors to be categorized into homogeneous cluster, the regional saliency scores are calculated by proposed cost-function. As we have mentioned in the introduction part, many contrast based models were confronted with the limitations caused by contrast. Therefore we now define simple and effective objectness principles to detect multiple salient objects by considering the mentioned limitations.

- (1) The spatial variations of foregrounds are generally smaller than backgrounds, which are wildly scattered over the whole image.
- (2) A spatial densities of foreground on the image borders is smaller than the background.

Note that the above principles considered only spatial properties of object without contrast feature. Based on these principles, we proposed a cost function for determining saliency scores corresponding to each categorized regions. The proposed cost function is given as,

$$RS_i(v_i, b_i, \sigma) = \exp(-(v_i\sigma_1 \times b_i\sigma_2)) , i \in \{1, \dots, k\} \quad (4)$$

where v_i is the variation of pixel coordinates corresponding to each i -cluster, and b_i represents the overlap score between observed cluster and image borders, k the number of clusters calculated by clustering algorithms. We used $\sigma_1=0.4$ and $\sigma_2=0.6$, to weigh b more heavily than v . As a saliency score gets closer to 1, the likelihood of foreground is increased. The spatial variation with observed clusters on the image $I = \{C_1, C_2, \dots, C_k\}$ is determined as,

$$v_i(c_i) = \frac{1}{h_{c_i}^d} \sum_{(x,y) \in c_i} (x - \mu_x)(y - \mu_y) , i \in \{1, \dots, k\} \quad (5)$$

where the $\frac{1}{h_{c_i}^d}$ is the dimension of the observed cluster c_i , x and y are the spatial coordinates of the cluster. μ_x and μ_y the centroid of the observed cluster c_i . Where v has a normalized value $[0, 1]$. Then, we compute the spatial densities of the observed clusters in the image borders. The density score are calculated by,

$$b_i(c_i, B_r) = \frac{h_{c_i \in B_r}^d}{h_{B_r}^d}, \quad i \in \{1, \dots, k\} \tag{6}$$

where c_i denotes observed cluster, and $h_{B_r}^d$ denotes the dimension of image borders with range r (By analyzing experimental results, we have adopted a boundary range $r = 10$ pixels from the boundary region begins). b_i also has a normalized value from 0 to 1. The visualization result of the saliency score computation is shown in Fig. 5. Finally, all computed saliency scores are integrated for final salient detection result, and the computed saliency score maps are illustrated in Fig. 6 second row.

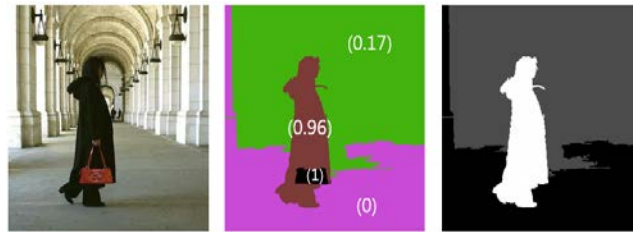


Fig. 5. Visualization result of the saliency score computation. From left to right : original image, categorized image (White numbers represent computed saliency scores corresponding to each region), result of the saliency score map.

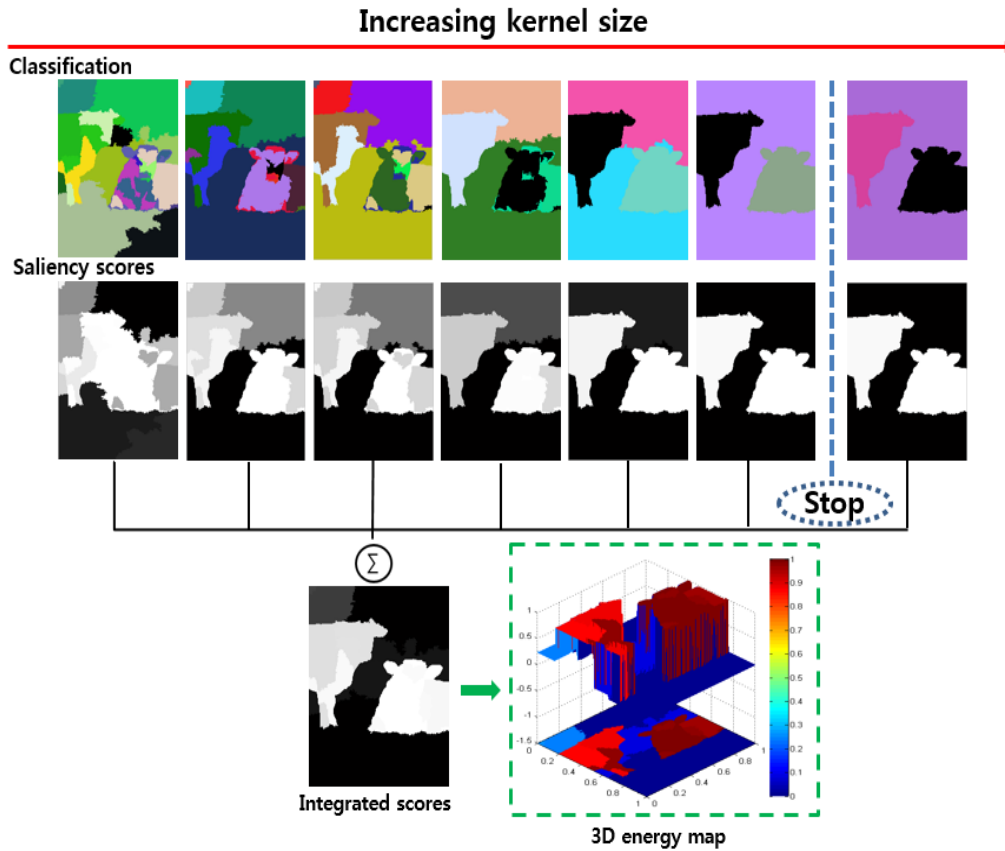


Fig. 6. Saliency score integration.

4. Experimental Results and Analysis

We have evaluated our method on three public benchmarks SED2 and MSRA-ASD, which represents multiple and single object images respectively, and compared it with eight state-of-the-art methods, FT [11], RC [21], DRFI [20], MZ [22], GB [13], IT [12], CW [16], and SF [18]. The main competitive methods are regional feature based methods [20], [21], which have achieved good performance compared to others. We selected competitive methods according to citations (IT, FT, GB and MZ), recency (DRFI, RC, SF and CW), better performance and being related to our work (DRFI and RC), and variety (IT is biologically-motivated, CW used the Gaussian weight map, FT works in frequency domain).

The saliency detection methods are evaluated using the precision-and-recall (PR) curve, which is commonly used for quantitative comparison of different models. In order to reliably compare various saliency detection methods, the salient regions should be evaluated by varying the careless threshold T from 0 to 255. Here precision represents the percentage of salient pixels corresponding to ground truth, whereas recall represents the ratio of salient pixel belonging to the total number of ground truth. The recall-and-precision metrics are computed at each binary image. The recall-and-precision metrics are computed as

$$recall = \frac{tp}{tp+fn} \quad , \quad precision = \frac{tp}{tp+fp} \quad (7)$$

To evaluate the performance of our proposed method, the average precision, recall, and F-Measure are compared over the entire ground truth database, with F-measure (we used $\beta^2 = 0.3$) defined as

$$F_{\beta} = \frac{(1+\beta^2)Precision \times Recall}{\beta^2 Precision + Recall} \quad (8)$$

The experiments have been tested on an Intel(R) Core(TM) i5-4670 CPU 3.40GHz and **Table 2** compares the average computational time of our models with competitive methods. The source codes of all competitive methods are provided by authors. **Table 2** shows that some methods evaluated in a C++ environment are faster than those implemented by MATLAB.

Table 2. Comparison of average time

Method	Type	Time(s)
FT [11]	MATLAB	0.15
IT [12]	MATLAB	0.67
DRFI [20]	MATLAB	8.48
MZ [22]	C++	0.27
CW [16]	MATLAB	0.78
SF [18]	MATLAB	0.20
RC [21]	C++	0.25
Ours	MATLAB	1.23

4.1 Performance comparison in the multiple object dataset

We have tested our method on SED2 dataset containing multiple objects provided by [23]. The ground truth segmentation were obtained by asking human subjects to manually segment the gray scale images (the color source was also provided) into two or three classes with each image segmented by three different human subjects. Fig. 7 shows the precision vs recall curve and F-measure with state-of-art-works. The precision and recall curves clearly demonstrate that our approach based on the three proposed countermeasures outperforms other methods in detection of multiple objects. As can be seen from Fig. 8, we can identify our results along with existing methods. The limitation with contrast features is observed in the second, third and fourth rows (cows, ducks and buildings), and the limitation to miss small object is shown in the fifth row (balloons).

4.2 Performance comparison in the single object dataset

We also perform saliency detection on the MSRA-ASD dataset representing simple single object images which have 1,000 pixel-based ground truth indicated from the MSRA dataset. Note that ASD was made up of simple foreground and background images but it is the most commonly used dataset for evaluating performance of salient detection in the recent years. In Fig. 9, the RC [21] has achieved best performance, and precision-recall curves show that the curve of our work is very slightly lower than RC and DRFI but consistently higher than the rest. The reason for this result is that RC used the contrast value, which heavily depend on a highlight region therefore the method can focus on only a single object in the simple background image. The reason why we don't use the contrast is that the contrast is not outstanding in the multiple object images. As can be seen in first row of Fig. 10, we see that our method has detected two objects, however the ground-truth only represents a single object. This condition is one reason of decreasing our performance in the MSRA-ASD.

4.3 Proposed model analysis

In the proposed model, we have used 4 parameters: image border range r , weight values σ_1, σ_2 in the saliency score computation process, and the number of super-pixel regions k . By considering experimental results with Fig. 11, the parameters are set empirically as $r=10$, $\sigma_1=0.4$, $\sigma_2=0.6$ and $k=150$. The most benchmarks used in the experiments consist of large scale objects, so the initial number of super-pixels between 150 and 200 is empirically proper to detect objects. We also provide experimental results according to the combination of the proposed features; in Fig. 11 (d), the combination of the color, edge and spatial features has achieved the best performance while implementing our model. Note that there is no independent test about use of the spatial features, because the spatial features are just supportive part in the classification process. In Fig. 12, we can see an effect of the combined features. Firstly, the color features clearly preserve the object borders but an energy of its inner domains are poor (see Fig. 12 (b)). Secondly, the spatial features enhance inner domain of the objects, yet background noise is also increased at the same time (see Fig. 12 (e)). Finally, the result of the Fig. 12 (g) indicates that the edge features have reduced a noise at the background area. Note that the color features are fundamental component in the proposed model. Therefore, we cannot expect better results without the color features, and the results are presented in Fig. 12 (c, f).

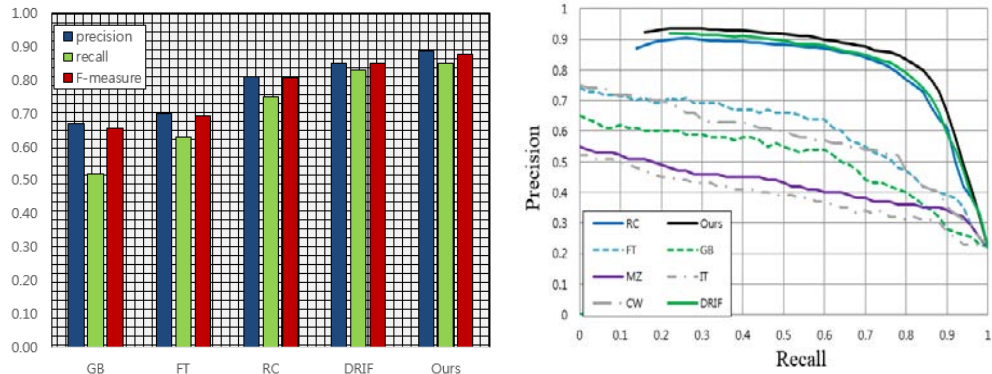


Fig. 7. Quantitative comparison of salient object detection results in the SED2. (a) Average precision, recall and F-measure (b) Precision-recall curve.

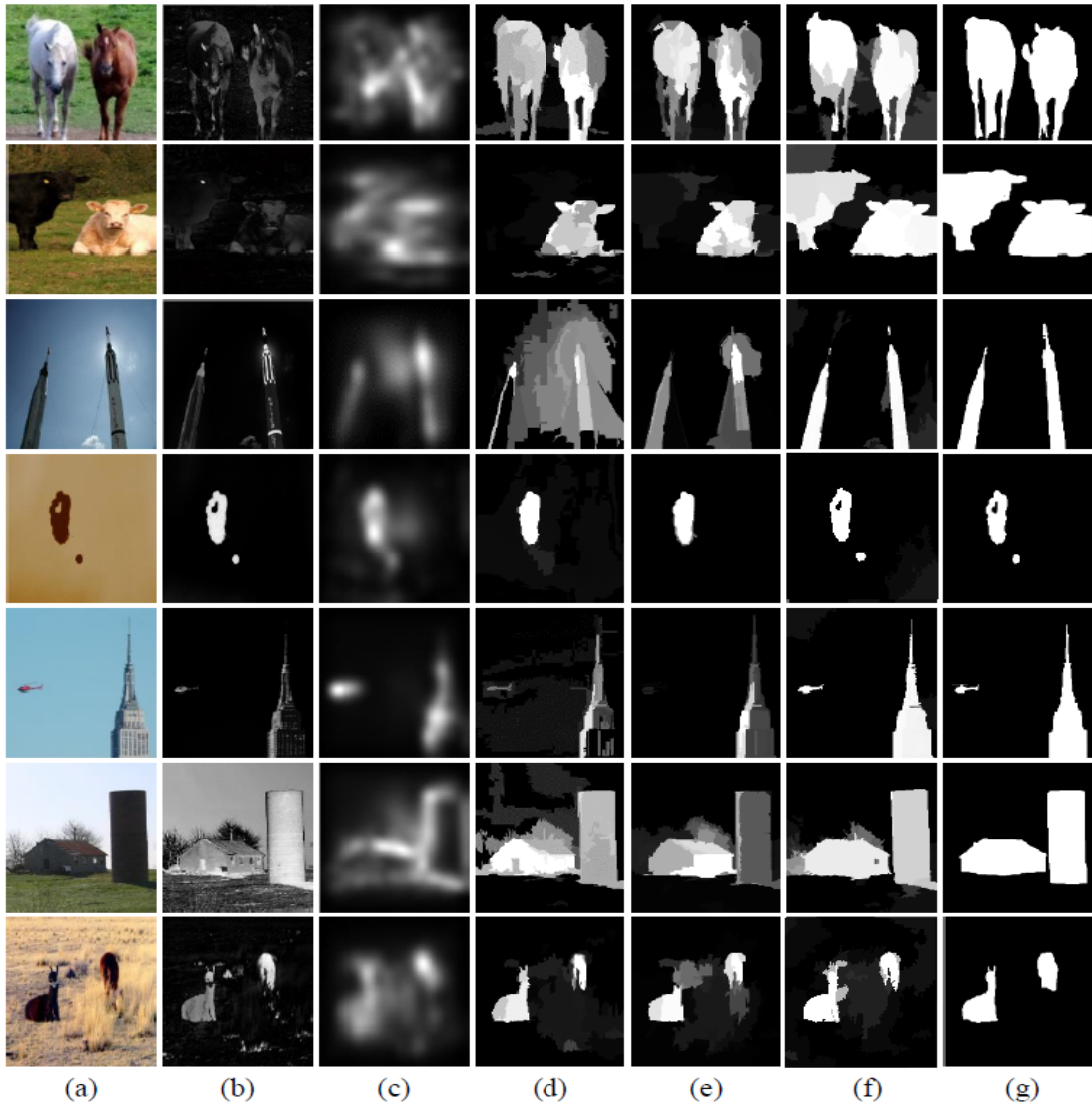


Fig. 8. Comparison of salient object detection results in the SED2. (a) Original (b) FT [12] (c) GB [24] (d) RC [20] (e) DRFI [21] (f) Our work (g) Ground-truth.

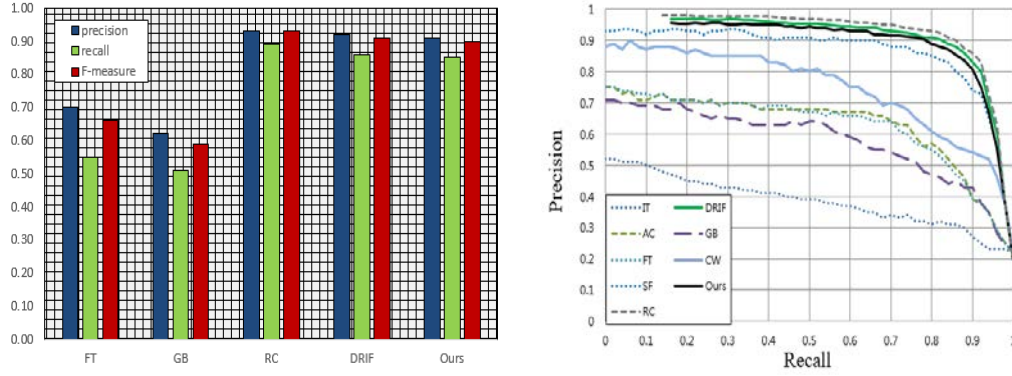


Fig. 9. Quantitative comparison of salient object detection results in the MSRA-ASD. (a) Average precision, recall and F-measure (b) Precision-recall curve.

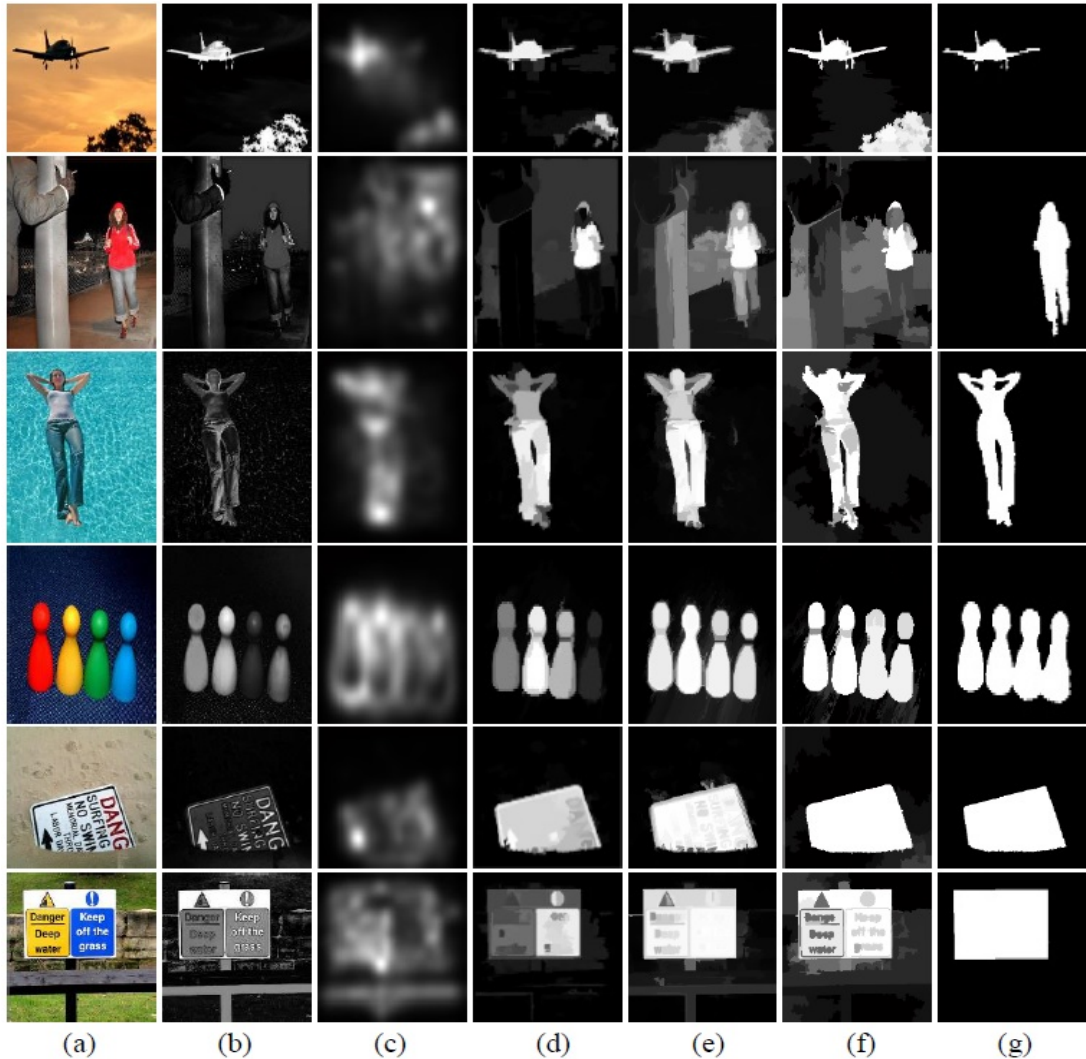


Fig. 10. Comparison of salient object detection results in the MSRA-ASD. (a) Original (b) FT [12] (c) GB [24] (d) RC [20] (e) DRFI [21] (f) Our work (g) Ground-truth

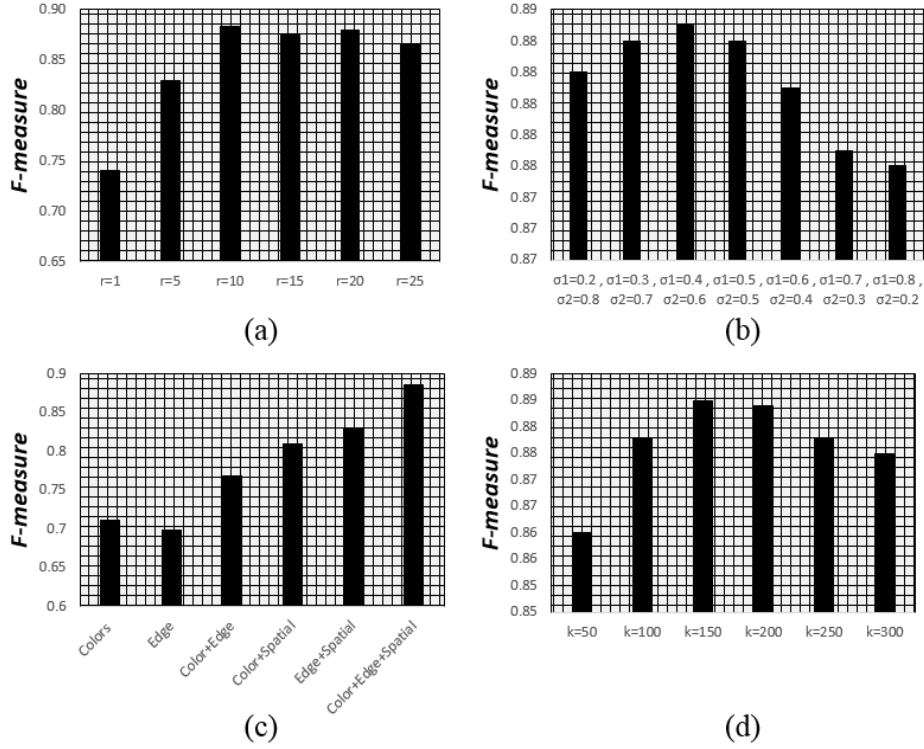


Fig. 11. Comparison of the proposed model with different parameters. (a) Image border range r and weight values σ_1, σ_2 , (b) Combination of the used features (d) Super-pixel initial parameter k .

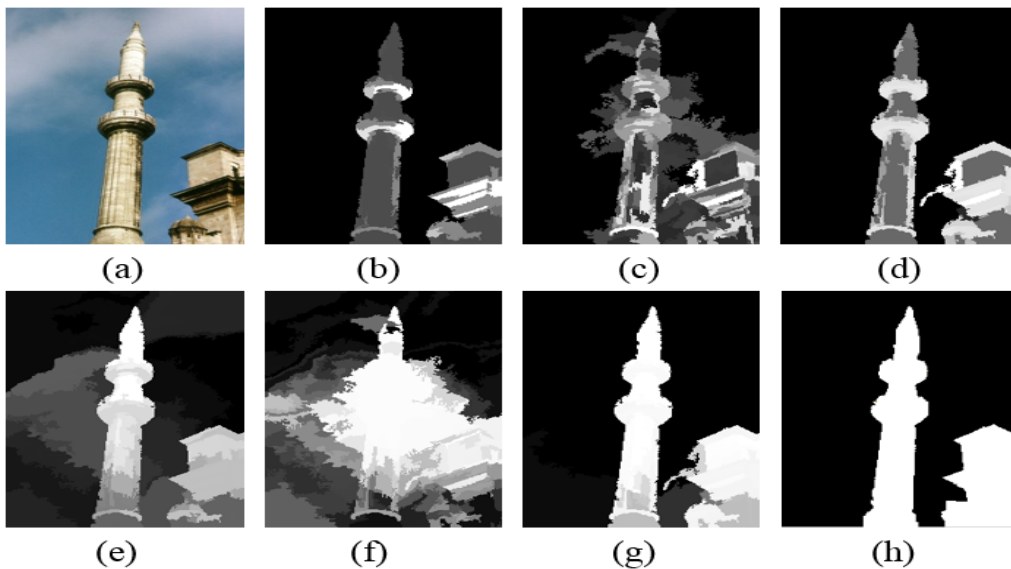


Fig. 12. Visualization results from the combination of the proposed features. (a) Original image (b) Color (c) Edge (d) Color + Edge (e) Color + Spatial (f) Edge + Spatial (g) Color + Edge + Spatial (h) Ground-truth.

5. Conclusion

In this paper, we address the salient region detection for multiple object images. By analyzing limitations of existing methods with multiple object images, we identify problems of contrast feature and introduce new objectness descriptor based on spatial feature without using the contrast features. We evaluated our method with eight state-of-art saliency detection models on SED2 and MSRA-ASD benchmarks. The experimental results demonstrate that our method has achieved superior performance in terms of both precision and recall on the multiple object images. However, the contrast features, which have achieved the best performance in the MSRA-ASD dataset, are still reasonable attempt to detect salient region in the single object images. In the future, we plan to continue to enhance the performance of the proposed method in the single object benchmark.

References

- [1] A. Berengolts and M. Lindenbaum, "On the distribution of saliency," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 1973-1990, 2006. [Article \(CrossRef Link\)](#)
- [2] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Human Neuro.*, vol. 4, no. 4, pp.219-227, 1985. [Article \(CrossRef Link\)](#)
- [3] L. Ye, S and Yuan Jun, "Salient object detection in video by optimal spatio-temporal path discovery," in *Proc. of the ACM MM*, pp.509-512, 2013. [Article \(CrossRef Link\)](#)
- [4] P. Khuwuthyakorn, A. Robles-Kelly and J. Zhou, "Object of interest detection by Saliency Learning," in *Proc. of the ECCV*, pp. 636-649, 2010. [Article \(CrossRef Link\)](#)
- [5] J. Bonaiuto and L. Itti, "Combining attention and recognition for rapid scene analysis," in *Proc. of the CVPR*, pp. 90-97, 2005. [Article \(CrossRef Link\)](#)
- [6] Z. Li, S.Qin, and L. Itti, "Visual attention guided bit allocation in video compression," *Image. Video Comput.* vol. 29, no. 1, pp.1-14, 2011. [Article \(CrossRef Link\)](#)
- [7] R. Achanta and S. Susstrunk, "Saliency detection using maximum symmetric surround," in *Proc. of the ICIP*, pp.2653-2656, 2010. [Article \(CrossRef Link\)](#)
- [8] G. Kootstra, B. de Boer and L. Schomaker, "Predicting eye fixations on complex visual stimuli using local symmetry," *Cognitive Computat.*, vol. 3, no. 1, pp.223-240, 2011. [Article \(CrossRef Link\)](#)
- [9] P.L. Rosin, "A simple method for detecting salient regions," *Pattern Recognit.*, vol. 42, no. 11, pp.2363-2371, 2009. [Article \(CrossRef Link\)](#)
- [10] N. Bruce and J.K. Tsotsos, "Saliency based on information maximization," in *Proc. of the NIPS*, pp.155-162, 2005. [Article \(CrossRef Link\)](#)
- [11] R. Achanta, S. Hemami, F. Estrada and S.Susstrunk, "Frequency-tuned salient region detection," in *Proc. of the CVPR*, pp.1597-1604, 2009. [Article \(CrossRef Link\)](#)
- [12] S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp.315-326, 2007. [Article \(CrossRef Link\)](#)
- [13] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp.1254-1259, 1998. [Article \(CrossRef Link\)](#)
- [14] E. Rahtu, J. Kannala and M. Salo, "Segmenting salient objects from images and videos," in *Proc. of the ECCV*, pp.366-379, 2010. [Article \(CrossRef Link\)](#)
- [15] D. Klein and S. Frintrap, "Center-surround divergence of features tatistics for salient object detection," in *Proc. of the ICCV*, pp.2214-2219, 2011. [Article \(CrossRef Link\)](#)
- [16] O. Kang-Han , K. Soo-Hyung, N. In-Seop, "Saliency detection using Centroid Weight Map," in *Proc. of the ACM ICUI MC*, Aticle No. 107, 2014. [Article \(CrossRef Link\)](#)
- [17] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no. 5, pp.603-619, 2002. [Article \(CrossRef Link\)](#)

- [18] Y.-F. Ma and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing," in *Proc. of the ACM MM*, pp.374-381, 2003. [Article \(CrossRef Link\)](#)
- [19] T. Liu, J. Sun, N.-N. Zheng and X. Tang, "Shum Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp.1-8, 2007. [Article \(CrossRef Link\)](#)
- [20] M.M. Cheng, G.X. Zhang, N.J. Mitra and X. Huang, S.M.Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp.569-582, 2015. [Article \(CrossRef Link\)](#)
- [21] J. Huaizu Jiang, W. Jingdong, Y. Zejian, W. Yang, Z. Nanning and L. Shipeng, "Salient Object Detection: A Discriminative Regional Feature Integration Approach," in *Proc of the CVPR*, pp 2083-2090, 2013. [Article \(CrossRef Link\)](#)
- [22] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc of the NIPS*, pp. 545-552, 2006. [Article \(CrossRef Link\)](#)
- [23] S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp.315-326, 2007. [Article \(CrossRef Link\)](#)
- [24] L. Ming-Yu, YL. Ming, T. Oncel, C. Rama, "Entropy Rate Superpixel Segmentation," in *Proc. of the CVPR*, pp. 2097-2104, 2011. [Article \(CrossRef Link\)](#)



Kang Han Oh received the B.S in Computer Science from Honam University in 2010. And he received the M.S in Electronic & Computer Engineering at Chonnam National University in 2013. He has been taking the Ph.D course in Electronics & Computer Engineering at Chonnam National University, Korea. His research interests are object detection, document image processing, text localization, medical image processing and pattern recognition.



Soo Hyung Kim received his B.S degree in Computer Engineering from Seoul National University in 1986, and his M.S and Ph.D degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1988 and 1993 respectively. From 1990 to 1996, he was a senior member of research staff in Multimedia Research Center of Samsung Electronics Co., Korea. Since 1997, he has been a professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing, and ubiquitous computing.



Young Chul Kim received his PhD from Michigan State University, USA, the MS from the University of Detroit, USA, and BS in electronics engineering from Hanyang University, Korea. In 1993, he joined the Department of Electronics Engineering at Chonnam National University (CNU) where he is currently a professor. From 2000 to 2004, he was a director of IDEC at CNU. From 2004 to 2005, he was a Vice Dean of the College of Engineering in this university. From 2004 to 2014, he was the chief of the LG Innotek R&D center at CNU. His research interests are SoC design and smart interface and Natural User Interface(NUI) and low power design.



Yu Ra Lee received her B.S. degree in Software Engineering from Chonnam National University, Korean in 2015. Since then, she has been a master process student in the Department of Computer Science, Chonnam National University, Korea. Her research interests are pattern recognition, image processing, and EPS gesture recognition.