

Salient Region Extraction based on Global Contrast Enhancement and Saliency Cut for Image Information Recognition of the Visually Impaired

Hongchan Yoon¹, Baek-Hyun Kim², Mukhiddinov Mukhriddin³, Jinsoo Cho⁴

¹Software Development Division, PCT Inc.
Seongnam, 13174, Korea
[e-mail: yhc@powerct.kr]

²KRRI(Korea Railroad Research Institute)
[e-mail: bhkim@krri.re.kr]

³Department of IT Convergence Engineering, Gachon University
Seongnam, 13120, Korea
[e-mail: mukhiddinovm@gc.gachon.ac.kr]

⁴Department of Computer Engineering, College of IT, Gachon University
Seongnam, 13120, Korea
[e-mail: jscho@gachon.ac.kr]

*Corresponding author: Jinsoo Cho

Received June 21, 2017; revised November 7, 2017; revised December 22, 2017; accepted January 18, 2018; published May 31, 2018

Abstract

Extracting key visual information from images containing natural scene is a challenging task and an important step for the visually impaired to recognize information based on tactile graphics. In this study, a novel method is proposed for extracting salient regions based on global contrast enhancement and saliency cuts in order to improve the process of recognizing images for the visually impaired. To accomplish this, an image enhancement technique is applied to natural scene images, and a saliency map is acquired to measure the color contrast of homogeneous regions against other areas of the image. The saliency maps also help automatic salient region extraction, referred to as saliency cuts, and assist in obtaining a binary mask of high quality. Finally, outer boundaries and inner edges are detected in images with natural scene to identify edges that are visually significant. Experimental results indicate that the method we propose in this paper extracts salient objects effectively and achieves remarkable performance compared to conventional methods. Our method offers benefits in extracting salient objects and generating simple but

A preliminary version of this paper appeared in the 2nd International Conf. on Electronics, Electrical Engineering, Computer Science: Innovation and Convergence (EEECS), August 10-13, 2016, Qingdao, China.

important edges from images containing natural scene and for providing information to the visually impaired.

Keywords: Salient region extraction, Visual attention, The visually impaired, Saliency map, Saliency cut, Image enhancement

1. Introduction

The human visual system (HVS) has a remarkable ability to quickly grasp the most salient region, allowing us to easily process contextual information of a scene based on the selective visual attention mechanism in an efficient manner. Visual saliency is the perceptual quality that makes certain objects stand out from the surrounding regions and capture human visual attention [1]. Humans are capable of identifying the most visually significant foreground object out of a scene, referred to as the salient object, in a quick and accurate manner, and adaptively focusing our attention on the perceived important region [2]. Attention is the essential link between cognition and perception. The ability to control the focus of attention can be driven either by stimulus (bottom-up mechanism) or goal (top-down mechanism) in human perception [3]. For this reason, saliency detection has been extensively studied by researchers in psychology, cognitive neuroscience, and computer vision. Methods for detecting salient region can be categorized into two primary approaches: bottom-up stimuli-driven and top-down goal-driven. Salient objects usually show a significant amount of contrast against their surroundings and are typically more compact in the image domain. Saliency algorithms based on these two cues usually work well if the contrast and diversity of the salient object are distinctive. However, saliency algorithms tend to fail in processing images containing a complicated background with various patterns [4]. The region of interest that attracts our attention in a bottom-up approach must be sufficiently distinctive with respect to surrounding features [5]. Bottom-up attention is fast, involuntary, feed-forward, stems from visual distinctness, uniqueness, and unpredictability. At the same time, it is closely related to image characteristics such as color, gradient, edges, and boundaries. In comparison, top-down attention tends to be slow, task-driven, voluntary, and closed-loop. Recently, visual attention modeling has become a popular research topic. It has been widely used in various computer graphics and vision techniques, such as object segmentation [6, 7], image retrieval [8, 9], smart video presentation [10], adaptive image compression [11], image retargeting [12], object recognition [13], sensation enhancement [14], content-aware image editing [15, 16, 17], video browsing [18, 19], and video saliency [20]. Even though significant progress has been made in the field of visual attention, it remains one of the most challenging topics in image analysis, pattern recognition, and computer vision.

Salient object detection is closely related to selective processing in the human visual system and aims to locate important objects in images [21]. Therefore, extraction of salient objects from images containing natural scenes can assist the visually impaired to identify

important information contained in images based on the readable and understandable tactile representation of salient objects. Generally speaking, tactile representation of salient objects in images should be designed with simplicity, so that they can be easily understood by the visually impaired. Hence, finding an optimum method for extracting salient objects from images is essential.

In this paper, a novel method is proposed for extracting salient regions based on global contrast enhancement and saliency cuts, in order to improve the image recognition process of the visually impaired. First, a global color contrast enhanced image is developed to improve the quality and contrast of the original image. Second, we use a saliency detection model based on global contrast to obtain a saliency map. Then, a binary saliency mask is obtained using adaptively triple thresholding along with the Graph Cut algorithm. The proposed method highlights salient objects, suppresses background, and extracts salient objects' outer contours and inner edges in order to help the visually impaired to better understand the content of an image scene.

Main contributions of the proposed method are as follows:

- Generating an enhanced image with global color contrast based on global histogram equalization.
- Salient region extraction from a saliency map using adaptive triple thresholding along with the Graph Cut algorithm.
- Extracted salient region translated into an assistive technology system, which is regarded as the next-generation technology of graphic haptic electronic board [22], for the visually impaired, as tactile graphics allow them to recognize and understand natural scene images.

This paper is structured as follows: Section 2 introduces conventional methods for detecting salient regions and saliency cuts, as well as for generating tactile graphics. Section 3 presents an overview of the proposed method. Section 4 explains the details of the proposed method and provides a summary of the proposed algorithm. Section 5 demonstrates the experimental results based on widely used benchmark databases. Finally, section 6 provides the conclusions of the study.

2. Related Work

In this section, we broadly review previous academic work related to several research areas: salient region detection, saliency cuts, saliency detection for the visually impaired and tactile graphic generation. We briefly look at existing saliency methods based on aforementioned research areas in the subsections to follow.

2.1 Salient Region Detection

In computer vision, current saliency detection algorithms can be classified into two categories: fixation prediction and salient object detection. The eye fixation prediction approach involves physically acquiring continuous movement of human attention based on

eye tracking, while the aim of the salient object-based approach is to identify salient objects from the input image. Saliency object detection is further extended to “objectness estimation” in object recognition. [23] Both are important and benefit different applications in high-level scene analysis [24]. Saliency models based on bottom-up methods convert natural scene images into saliency maps, where each pixel/superpixel and region is assigned a saliency value or probability.

Bottom-up approaches can be divided into two groups: statistical and spectral methods. Statistical methods are based on the center-surround hypothesis. The early models were designed to predict human eye fixations and were evaluated according to datasets by introducing the principle of saliency detection. Ullman and Koch [25] define saliency as how different an image is from its surroundings in color, orientation, motion, depth etc. Itti et al [26] follow the same framework and propose a biologically inspired center-surround contrast based on Difference of Gaussian to acquire the saliency map. By extracting primary visual features, three visual channels are calculated in parallel and independently. Harel et al. [27] propose graph-based visual saliency using pixel positions and weight values that are proportional to the dissimilarity obtained from intensity, orientation, and their variations. Achanta et al. [28] propose a frequency tuned approach, which involves acquiring a saliency map from the color differences of the entire image after blurring is used in the surrounding region. Ma and Zang [3] use similar color contrast analysis to compute the center-surrounding color difference in a fixed neighborhood for each pixel. Mengnan Du et al. [29] transform an image into a perceptually homogeneous region in order to calculate three region contrast measures: global distinctness, region compactness, and center. This is performed prior to obtaining an initial saliency value of each image region. Liu et al. [30] demonstrate a set of novel features – multi-scale contrast, center-surround histogram, and color spatial distribution – to designate a salient object locally, regionally, and globally. The concept of Conditional Random Field is introduced to effectively combine these features for salient object detection.

In addition, pixel/superpixel and region-based methods have been explored. These methods first apply image segmentation techniques (e.g., graph-based [31], mean shift [32], and superpixels [33]) to the input image and segment homogeneous regions to extract feature statistics from each segmented region for salient region detection. Cheng et al [2] demonstrate two saliency models: histogram-based contrast (HC), which designates a pixel-wise saliency value, and region-based contrast (RC), which reflects spatial relations at the cost of reduced computational efficiency. Inspired by Cheng’s method [2], Lou et al. [34] propose a regional principal color contrast saliency detection modal, which integrates low-level and medium-level visual cues. They initially detect color saliency based on global regard, then a local model segments the quantized input image into regions using graph-based image segmentation. Lastly, spatial relationships are utilized to generate a final saliency map. Perazzi et al. [35] generate saliency by decomposing an image into compact and perceptually homogeneous elements that abstract unnecessary details. Hua et al. [36] introduce color contrast using background priors, which allow colors in the image boundary to represent all background colors. They first smooth the image and then calculate the saliency of the image boundary and separate salient patches from the

background using adaptive thresholding. Han et al. [37] develop a framework for saliency detection by first modeling the background and then separating salient objects from the background. Color contrast and spatial distribution are used to obtain pixel-accurate saliency maps. Jiang et al. [38] propose region-level saliency detection using both global and local color contrast to measure the spatial compactness. They use the SLIC algorithm to segment input image into superpixels and calculate the quantized color histogram. They then compute color contrast and similarity distribution, and combine them with central bias prior. Zhou et al. [39] combine widely used contrast measurements such as center-surround, corner-surround, and global contrast to detect visual saliency. Zhou et al. [40] propose an adaptive-based region merging algorithm to reassemble regions based on color dissimilarities. They also use the SLIC algorithm for initial segmentation and the superpixels located on the image boundary are treated as seeds for initiating the merging process. Qian et al. [41] propose a saliency detection framework using the new concept of optimal contrast. First, they utilize the algorithm of weighted sparse coding residual (WSCR) to calculate contrast. Afterwards, they use an entropy-based scheme to select the optimal contrast from contrast hypotheses, and the optimal contrast is used to evaluate visual saliency.

Another popular approach is the spectral method. These methods endeavor to decompose the background efficiently by analyzing filter responses in the frequency domain. This is based on the assumption that less periodicity represents a rare event (i.e., saliency) on the corresponding location. Zhang et al. [42] introduce global contrast in the frequency domain by analyzing the log-spectrum of an image, which transforms the spectral residual into the spatial domain to obtain a saliency map. Guo et al. [43] propose a quaternion representation of an image composed of intensity, color, and motion features to build a spatiotemporal saliency map. The added dimension of motion allows the phase spectrum to represent spatiotemporal saliency so that the level of attention can be selected for images and videos.

2.2 Saliency Cuts

Approaches based on saliency cuts have been widely explored in recent years because these methods provide a binary mask of salient objects assisted by a saliency map. Saliency cuts automatically segment an object from its background. With a saliency map and using the iterative GrabCut algorithm [44], we can extract a precise image mask [2, 45, 46, 47]. Fu et al. [47] modify the graph cut method by examining the effects of labels in graph-based segmentation. Aytakin et al. [48] propose a link between quantum mechanics and spectral graph clustering, referred to as Quantum Cuts, which produces a graph among superpixels extracted from an image, and optimizes a measure related to the image boundary, local contrast, and area data. Peng et al. [49] demonstrate a saliency-aware stereo image segmentation approach using the disparity map and statistical information of stereo images to enrich high-order potentials. Mehrani et al. [46] utilize the standard features used in vision based on color and texture. Properly normalized, these simple features yield performances superior to methods based on hand-crafted characteristics that are specifically designed for detecting saliency.

2.3 Saliency Region Detection for the Visually Impaired

Observing, recognizing, and understanding visual information have always been highly significant issues for the visually impaired. Although saliency detection is utilized in obstacle detection and environmental awareness applications for the visually impaired, generating meaningful visual data across images remains a challenging problem. Chen et al. [50] propose a system for detecting obstacles based on visual selective features and stereo vision. By eliciting feature vectors containing color, direction, and intensity, a corresponding saliency map is constructed. Then, an obstacle on the road can be located after thresholding of the feature map, and the 3D data of the obstacle location can be calculated using stereo vision. Lastly, the information about the obstacle is transformed into text, and the voice information is delivered to the user. Muthulakshmi et al. [51] propose a method for tracking and recognizing an object for the visually impaired. Their system is based on Ullman and Koch's [25] hypothesis, which is an early model of eye fixation. Wang et al. [52] introduce a navigation and wayfinding application that detects doors and signage to help the visually impaired to navigate unknown indoor environments. They detect indoor signage and extract attended areas using a saliency map and employ bipartite graphic matching to recognize indoor signage in only important areas rather than the entire image. For extracting a salient region, Wang et al. compute saliency maps using color, intensity, and orientation for each pixel based on a linear "center-surrounded" operation.

2.4 Tactile Graphic Generation

Generating tactile graphics is one of the most challenging research topics because the techniques involve obtaining critical information from a given image. Takagi et al. [53] propose an algorithm for extracting character strings from scene images based on edge detection, morphology operator, and fuzzy inference. This framework allows the visually impaired to walk more independently on the street. Jungil et al. [22] propose an education assistive technology system based on a graphic haptic electronic board. This system enables authoring of education materials, as well as automatic conversion and real-time distribution of academic materials for low-vision and blind students. Chen et al. [54] present a method of recognizing mathematical graphs and for extracting and classifying broken line elements from mathematical graphs by utilizing techniques introduced for developing a computer-aided system to extract and classify broken lines. Chen et al. [55] also propose a method for automatically translating hand-drawn maps into tactile maps using a pattern recognition technique for extracting and classifying objects in hand-drawn maps.

3. Overview

In this section, we provide a brief overview of the proposed salient region extraction method. It consists of four steps: 1) enhancing global contrast, 2) generating a saliency map,

3) extracting a salient region using saliency cuts, and 4) generating salient object's boundaries and inner edges. As shown in Fig. 1, we first use an image enhancement technique to enhance the image quality and increase the color contrast of significant objects against the surrounding regions. Second, we divide the image with enhanced contrast into several regions using saliency values, and then produce a saliency map based on a conventional saliency method [2]. Third, we obtain a salient region to produce a binary mask of high quality using saliency cuts [45]. Finally, we extract a salient object in a full color space and extract salient object's contours and inner edges by removing noise and unnecessary details. We then convert the boundaries and edges extracted from the procedures into tactile graphics. Fig. 2 shows the four steps involved in the proposed method.

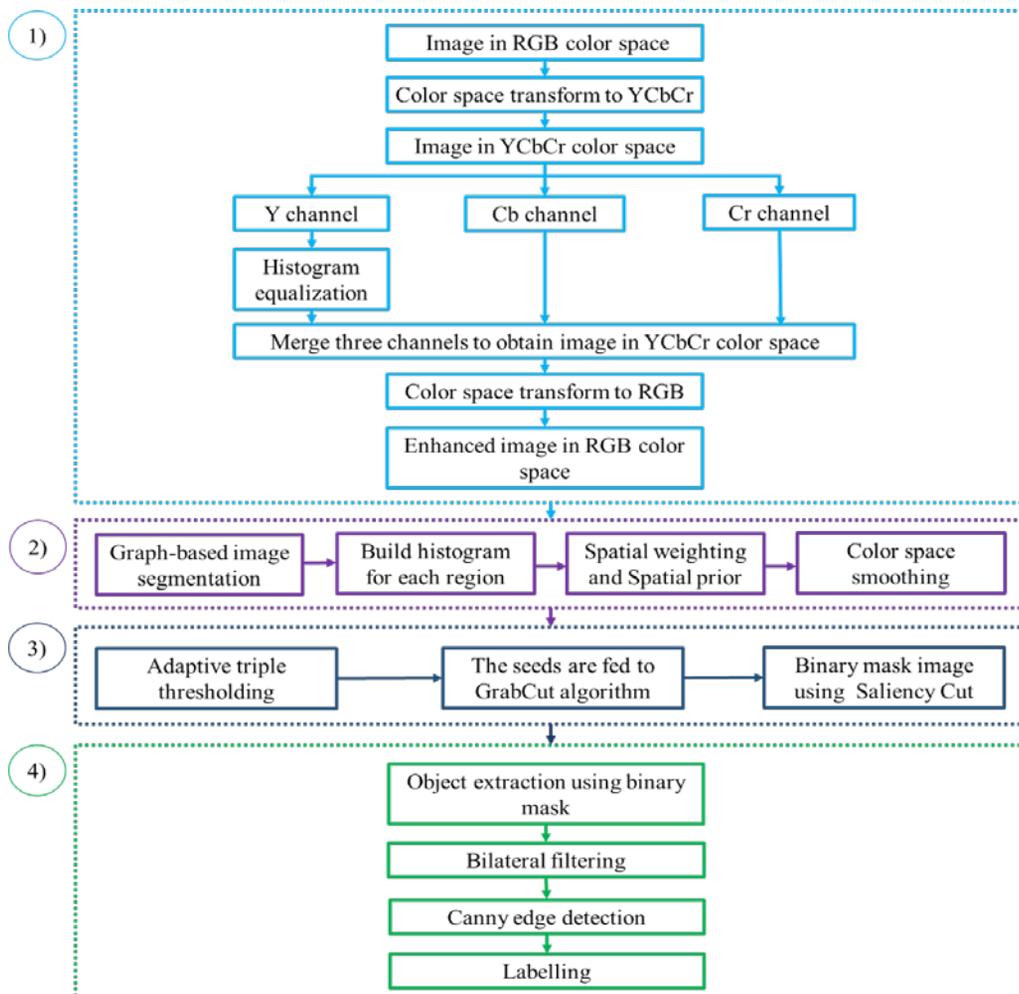


Fig. 1. Overall process of the proposed method: 1) enhancement of global contrast, 2) generation of saliency map, 3) extraction of salient region using saliency cuts, and 4) generation of salient object's contour and inner edges.

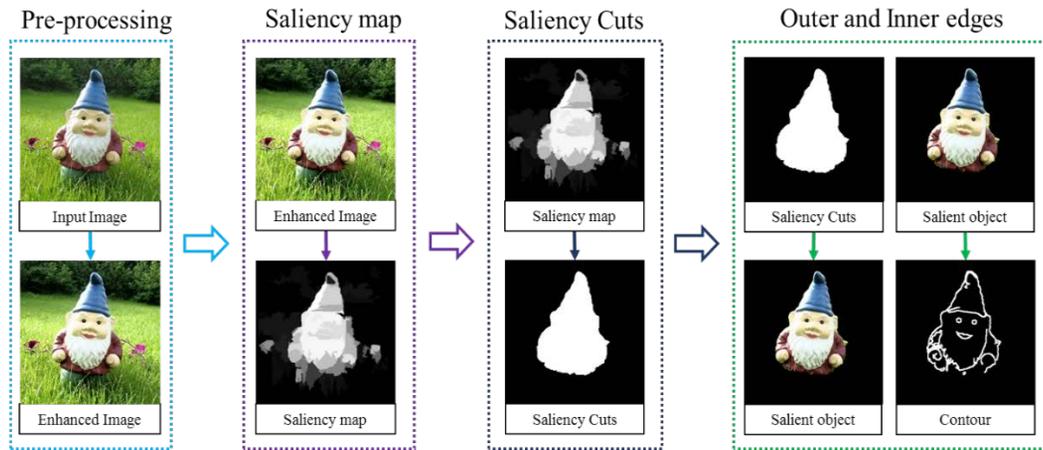


Fig. 2. Four steps in the proposed method: enhancement of global contrast, generation of saliency map, extraction of salient region using saliency cuts, and generation of salient object's boundary and inner edges.

4. Algorithms

This section describes the details of the proposed method. We first introduce a global contrast enhancement technique in subsection 4.1. Then, we use color contrast enhanced image to obtain a saliency map as explained in subsection 4.2. Subsection 4.3 defines a high quality binary salient mask created using saliency cuts based on adaptive triple thresholding. Lastly, subsection 4.4 determines the outer and inner edges of a salient object by employing post-processing techniques such as bilateral filtering and Canny edge detection.

4.1 Global contrast enhancement

Human biological vision system is very sensitive to contrast in visual information. Therefore, we use an image enhancement technique to increase the contrast of a given image. The objective of enhancing global color contrast is to provide a clear image to facilitate extracting regions or objects from a natural scene image. One of the simplest and widely used methods for enhancing contrast is histogram equalization (HE). The principal purpose of using the RGB color space is to display images in an electronic system. However, it is better to use other color spaces (LAB, YCbCr, and HSV) in some applications. An RGB image is color-coded as 24 bits per pixel, and the RGB color space does not have any inter-channel arithmetical operation, as opposed to other color spaces such as LAB, YCbCr, and HSV. The RGB format cannot be used in our study because all of its three planes represent color components red, green, and blue. Since the histogram of the intensity component (rather than the color component) has to be equalized, we need to convert the RGB color space into the YCbCr space (Y: luminance, Cb: chrominance-blue,

Cr: chrominance-red), which is used in NTSC and PAL TV systems. First, each channel of the multi-channel array is split into separate channels, and only the first luminance channel (Y) is used for histogram equalization. It is worth noting that other color channels are unchanged and merged with the color components with processed luminance. In YCbCr, each channel has 8-bit encoding.

- Y ranges from 16 to 235, 16 being the darkest and 235 the brightest.
- Cb and Cr range from 16 to 240, with 128 being the reference point.

The following relationships are used to transform pixels between RGB and YCbCr color spaces.

$$Y \leftarrow 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

$$Cr \leftarrow (R - Y) \cdot 0.713 + \text{delta} \quad (1)$$

$$Cb \leftarrow (B - Y) \cdot 0.564 + \text{delta},$$

where $\text{delta} = 128$ for an 8-bit image. Y, Cr, and Cb cover the entire value range. We convert the histogram-equalized image from YCbCr into RGB color space for further processing. We were able to obtain best results by improving image quality using global histogram equalization contrast to unsharp masking (UMS) as well as contrast limited adaptive histogram equalization (CLAHE). Therefore, we used global histogram equalization in the proposed method [56]. The sample results are shown in [Fig. 3](#).



Fig. 3. Results of global contrast enhancement. Top row: original images. Bottom row: images with enhanced contrast.

4.2 Saliency map generation

Image contrast and spatial relationships are important in human visual perception. A region with a high contrast against the surrounding region is typically a strong evidence of saliency. Approaches based on global contrast, which separate an entire salient object from its surroundings, are preferable to methods based on local contrast, which obtain saliency values at object boundaries. Therefore, a contrast enhanced image is segmented into regions using graph-based image segmentation, and then a saliency map is generated using salient region detection based on global contrast.

Saliency map methods based on global contrast calculate saliency of an image region based on its contrast against the entire image. We first segment a contrast enhanced image into regions. Felzenszwalb and Huttenlocher's segmentation method is used to generate an image region [31]. The steps for obtaining a saliency map is described in Cheng et al. [2]. Next, we build a color histogram for each region. For each region r_k , a saliency value is determined based on its color contrast against all other regions contained in the image.

$$S(r_k) = \sum_{r_i \neq r_k} w(r_i) D(r_k, r_i), \quad (2)$$

where $w(r_i)$ is the weight of region r_i , and $D(r_k, r_i)$ is the color distance metric between two regions. We obtain the color distance between two regions r_1 and r_2 as follows:

$$D_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_{1,i}) f(c_{2,j}) D(c_{1,i}, c_{2,j}), \quad (3)$$

where $f(c_{k,i})$ is the probability of the i -th color $c_{k,i}$ among all n_k colors in the k -th region r_k , $k = \{1, 2\}$. Spatial information is incorporated using a spatial weighting term to increase and decrease the effects of closer and farther regions, respectively.

$$S(r_k) = w_s(r_k) \sum_{r_i \neq r_k} e^{-\frac{D_s(r_k, r_i)}{\sigma_s^2}} w(r_i) D_r(r_k, r_i), \quad (4)$$

where $D_s(r_k, r_i)$ is the spatial distance between regions r_k and r_i , σ_s control the strength of spatial distance weighting, $w(r_i)$ is the weight of region r_i defined by the number of pixels in r_i , $w_s(r_k) = \exp(-9d_k^2)$ is a spatial prior weighting term, and d_k is the average distance between pixels in region r_k and the center of the image. We further improve the saliency

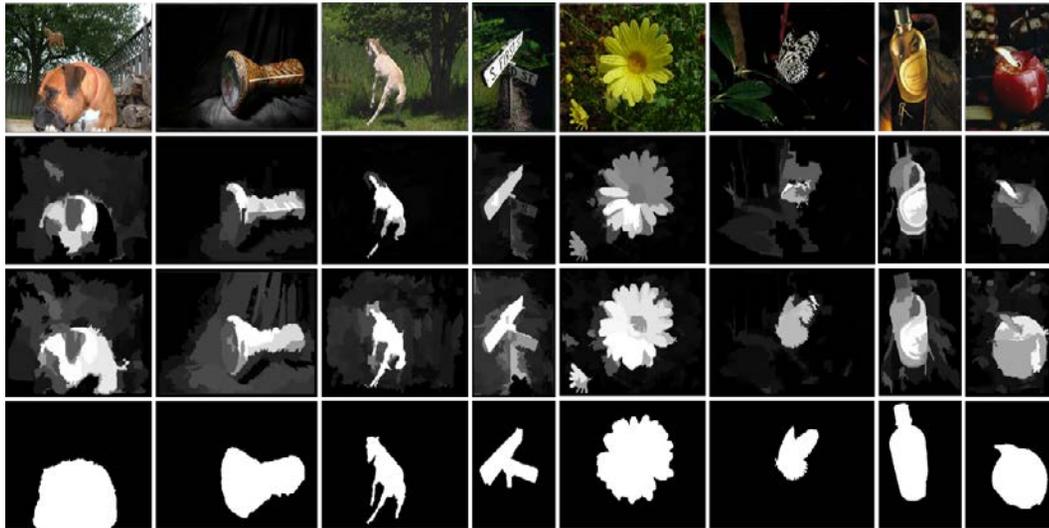


Fig. 4. Experimental results for detecting salient region. Top row: input images; second row: RC [2] method; third row: our method; and bottom row: ground truth images.

map using spatial prior and by applying color space smoothing, as described in RC [2]. As shown in Fig. 4, the results of our salient region are better than RC in most cases, especially for images containing blurry and/or dark natural scene [2], thanks to pre-processing histogram equalization.

4.3 Saliency cuts for salient region extraction

The saliency maps generated from the previous procedures are used to assign an iterative version of GrabCut, referred to as saliency cut [45]. We used adaptive triple thresholding using the histogram of the saliency map. Three thresholds t_l , t_m , and t_h ($0 < t_l < t_m < t_h < H$) are calculated by optimally dividing the saliency histogram into four parts. We obtained the three-level thresholds using the Otsu algorithm [57], which is a popular binarization algorithm for its simplicity and good performance. The triple threshold values decompose the histogram of the saliency map into four parts: certain background T_{cb} , probable background T_{pb} , probable foreground T_{pf} , and certain foreground T_{cf} . These four seeds are fed into the GrabCut algorithm, which generates a binary mask of a salient object. We obtain salient object's contour using the binary mask. Fig. 5 demonstrates how the outer boundary of a salient object is detected using Canny edge detection. These outer boundaries can be converted into tactile graphics for the visually impaired. In some cases, however, the visually impaired may not be able to correctly identify an object just by touching its outer boundary. Therefore, we also propose a method for detecting inner edges of a salient object in natural scene images.

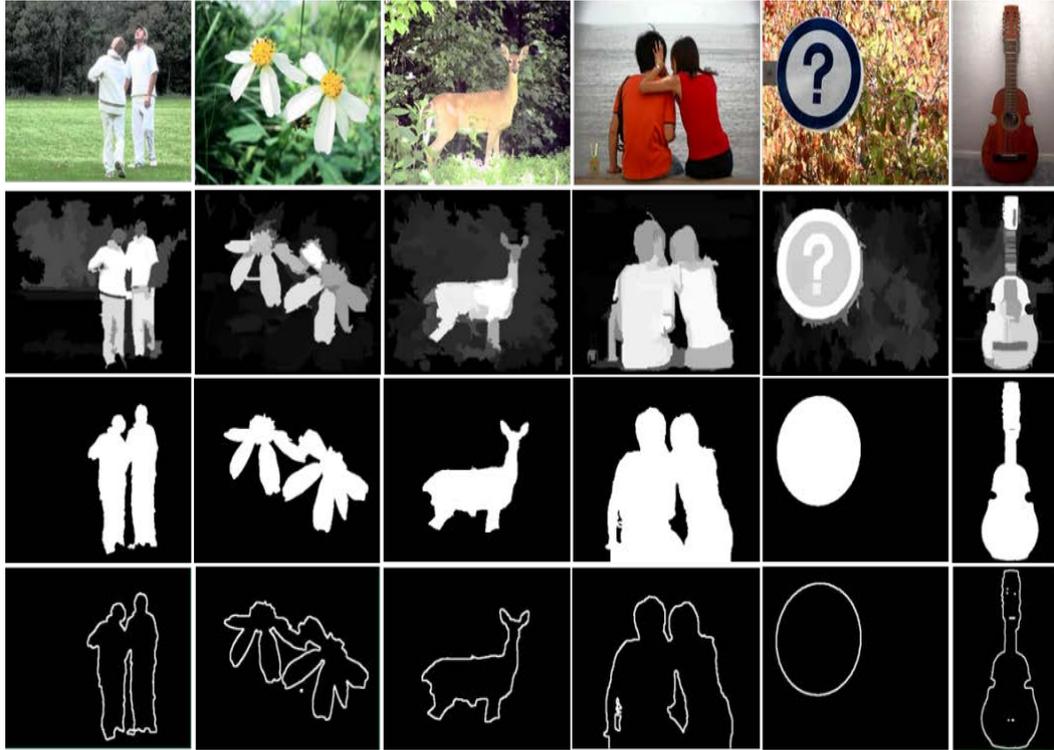


Fig. 5. Experimental results of outer contour generation. Top row: contrast enhanced image; second row: saliency map; third row: binary mask of salient region; and bottom row: Canny edge detection and salient object's outer contour.

4.4 Boundary and inner edge generation

Sometimes, the outer boundary of an object does not provide sufficient information for the visually impaired to correctly identify the salient object in a natural scene image. Therefore, we need a method for generating object's inner edges while highlighting the most important edges and suppressing noise and minute details. For this, we first obtain a salient object in a full color space using a binary mask of the salient object. We create a matrix the size and type of which are identical to the original input image, in order to calculate a desired output image. Then, we need to duplicate the non-zero elements of the binary mask that indicate the matrix elements of the original input image.

$$D_o = B_m(x, y) * I_i(x, y), \quad (5)$$

where D_o is the desired output image, $B_m(x, y)$ is the binary mask, and $I_i(x, y)$ is the input image. We can now obtain a salient object in a full color space, as shown in **Fig. 6**.

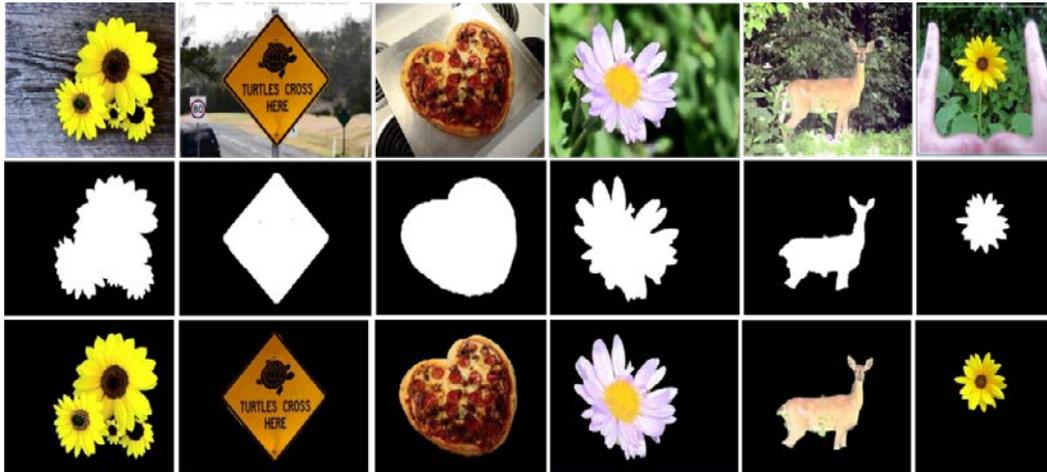


Fig. 6. Experiment results of the binary masking method. Top row: images with contrast enhancement; middle row: binary mask; and bottom row: salient object in a full color space.

Bilateral filtering [58] has long been an important technique in image processing. It is a simple, non-iterative scheme for smoothing images while carefully preserving edges. In addition, it can reduce unwanted noise while maintaining edges fairly sharp. This filtering method is used for edge detection. For simplicity and suitability for all kind of images, we assign an identical sigma value to the color space and the coordinate space. In our case, we use a sigma value of 150 because we need to reduce unwanted noises and produce a strong effect while maintaining edges fairly sharp. We set the filter size to 9. We obtain salient object's boundary and its inner edges using Canny edge detection [59], which is a widely used multi-step algorithm capable of detecting edges and suppressing noise simultaneously. (We have experimentally found that Canny edge detection performs better than many existing models, such as Sobel, Prewitt, and Robert.) Finally, we obtain salient object's contour and its inner edges with minimum image information in order for the visually impaired to easily understand the content of an image scene. Moreover, we can translate these contours and edges into tactile graphics for assistive technology systems.

5. Experiment and Evaluation

We used the C++ language to implement the proposed method and performed experiments on a PC with 2.60GHz CPU and 4GB of RAM. We also used an MSRA10k dataset [2], which includes 10,000 natural scene images and manually labelled masks. Both quantitative and qualitative comparisons of the proposed approach were performed with 12 alternate methods: region-based contrast (RC) [2], histogram-based contrast (HC) [2], context-aware saliency (CA) [60], appearance similarity and spatial distribution (GC) [61], contextual hypergraph modelling (CHM) [62], the absorbing Markov chain over an image

graph model (MC) [63], spatially weighted dissimilarity (SWD) [64], patch distinctness with an object probability map (MNP) [65], pattern distinctness using principal component analysis (PCA) [66], region covariance descriptors for visual saliency (COV) [67], highlighting sparse saliency regions (SS) [68], and looking beyond the image using its patches (LBI) [69].

5.1 Quantitative comparison

We use three standard measures that are widely recognized and easy to understand to compare techniques for detecting salient regions. We first use a precision and recall curve analysis for quantitative evaluation to obtain a receiver operating characteristics (ROC) curve. From these two metrics, we obtain the average F-Measure and the maximal F-Measure, which cooperatively reflect recall and precision, as well as the area under curve (AUC), which can be calculated according to the area under the ROC curve. Furthermore, we employ a third measure to measure the mean absolute error (MAE) between the evaluated saliency map and the ground truth annotation.

We utilize the measures used in [2, 61, 63], i.e., precision-recall (PR) curve, to evaluate the performance of the proposed method and those of alternative methods. Precision is the fraction of the detected salient pixel belonging to the salient object in the ground truth, and recall corresponds to the percentage of correctly assigned salient pixels. The PR curve is acquired by normalizing the saliency map to [0, 255] to generate binary masks with a threshold ranging from 0 to 255, and by comparing the binary mask with the ground truth image, which is manually labelled. Precision and recall rates are computed as follows:

$$precision = \frac{|TDO \cap GT|}{|ADO|}, \quad (6)$$

$$recall = \frac{|TDO \cap GT|}{|GT|}, \quad (7)$$

where TDO , ADO , and GT denote truly detected salient objects, all detected objects, and manually labelled ground truth, respectively. In addition, to compute the F-Measure value, which balances measurements between the mean of precision and recall rates and is proportional to performance. F-Measure is defined as follows:

$$F_{\beta} = \frac{(1 + \beta^2) precision \times recall}{\beta^2 \times precision + recall}, \quad (8)$$

where in order to assign more importance to the precision value, $\beta^2 = 0.3$ as suggested by a number of studies on salient object detection, including [2] and [45]. A higher weight is

assigned to precision than recall because recall can be easily acquired by establishing the entire region as a foreground. We used a fixed threshold varying from 0 to 255. For each threshold, a pair of precision and recall scores are computed and combined to form a precision-recall (PR) curve. The PR curve can be scored based on its maximal F-Measure as a good indication of the salient region detection performance when using fixed thresholding.

In addition to PR and F-Measure, we can define the false positive rate (FPR) and true positive rate (TPR) in order to binarize the saliency map using a series of fixed thresholds. The ROC curve can be obtained by plotting TPR against FPR and by varying the threshold. TPR and FPR are defined as follows:

$$TPR = \frac{|BM \cap GT|}{|GT|}, \quad (9)$$

$$FPR = \frac{|\overline{BM} \cap \overline{GT}|}{|\overline{GT}|}, \quad (10)$$

where \overline{BM} and \overline{GT} denote the opposite of the binary mask BM and the ground truth GT , respectively.

We also assess the area under ROC curve to calculate the AUC score. Although ROC is a two-dimensional (2D) representation of model's performance, the AUC process converts this information into a single scalar. As the name indicates, it is computed using the area under the ROC curve. An AUC score of 1 represents a perfect model, and arbitrary estimating will yield a score of around 0.5.

In some applications with weighted quality, a continuous saliency map may be more important than binary masks. Therefore, in order to perform a more comprehensive comparison, we also evaluate the MAE between the continuous saliency map SM and the GT , both normalized in the range of [0,1]. The MAE score is obtained as:

$$MAE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |SM(x, y) - GT(x, y)|, \quad (11)$$

where W and H denote saliency map's width and height, respectively, and $SM(x, y)$ and $GT(x, y)$ are the saliency value and the ground truth at the pixel position (x, y) , respectively.

Our method outperforms all other methods in terms of both PR and ROC curves based on an MSRA10k dataset and is comparable to RC [2] and MC [63]. The reason might be that the proposed approach uses global similarity and spatial distance, as in other two

saliency region detection methods. **Fig. 7** shows the PR and ROC curves of the proposed method along with other saliency methods.

Fig. 8 shows the F-Measure, AUC, and MAE scores. As shown in the left column, our method outperforms other methods, yielding higher mean F-Measure, max F-Measure, and AUC scores. As shown in the right column, the proposed method yields the lowest MAE score, which is inversely proportional to performance.

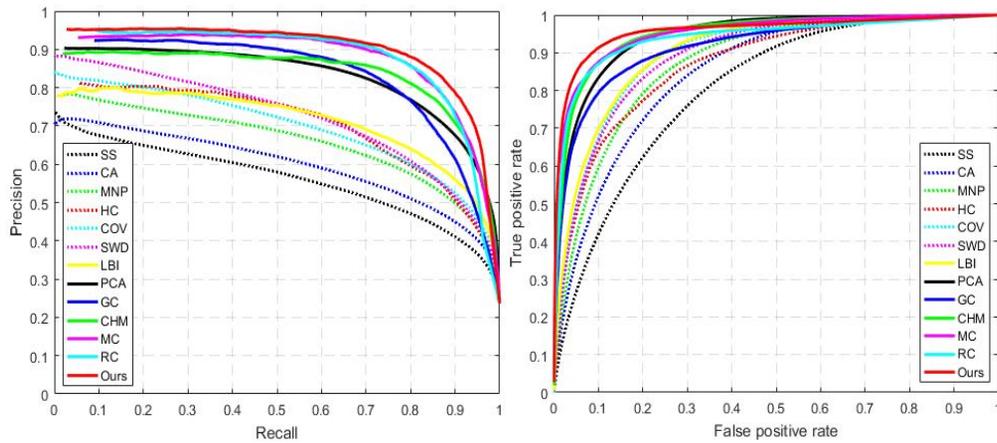


Fig. 7. Quantitative comparisons using the PR (left) and ROC (right) curves of saliency maps.

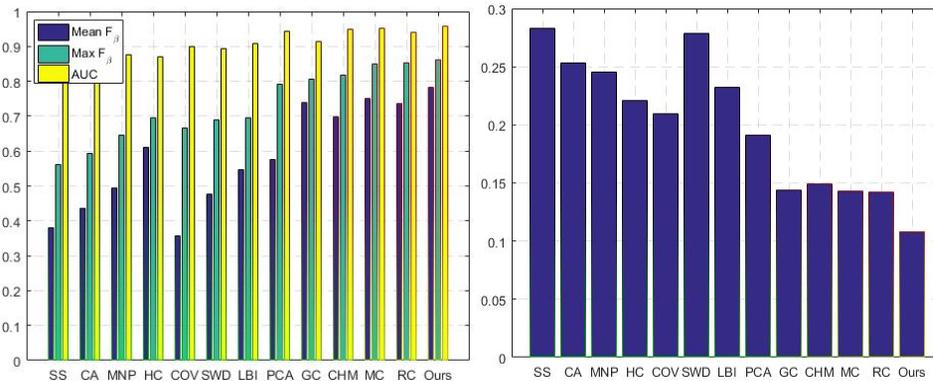


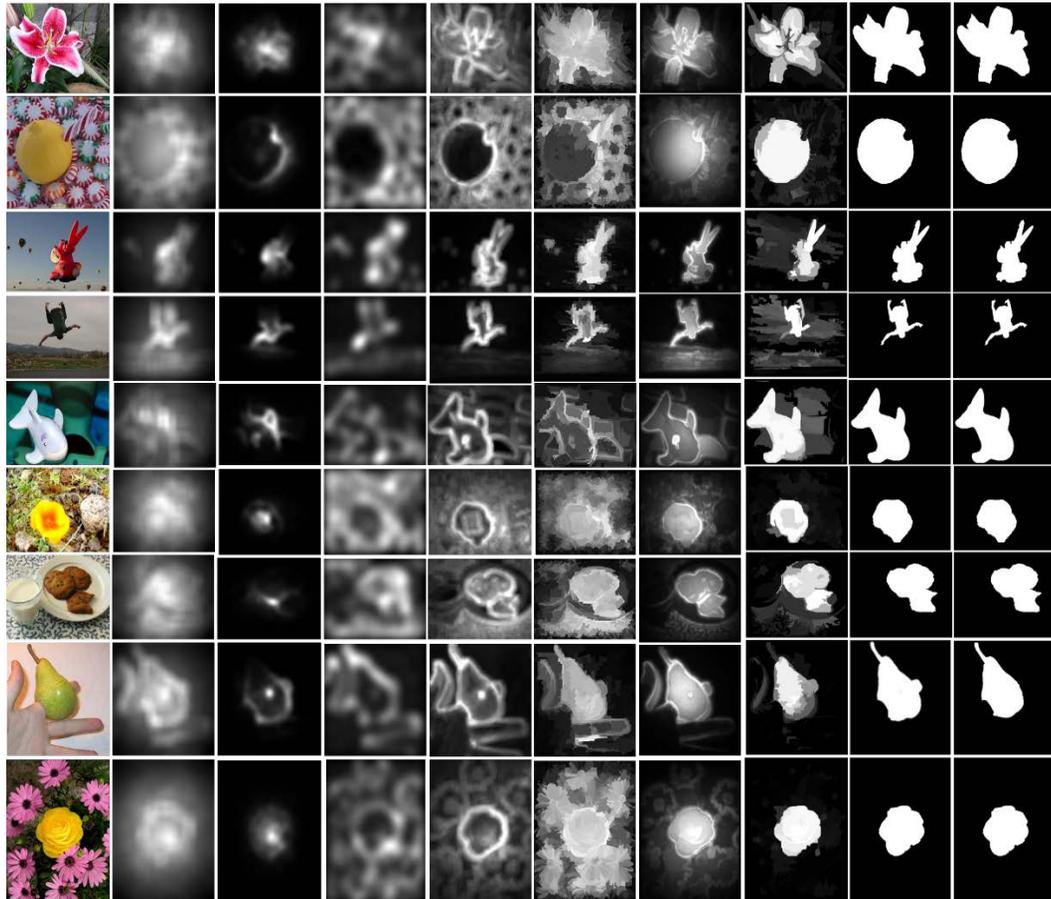
Fig. 8. Quantitative comparisons using the F-Measure (left), AUC (left) and MAE (right) scores of saliency maps. Abbreviations of other methods are defined as follows: highlighting sparse saliency regions (SS) [68], context-aware saliency (CA) [60], patch distinctness with an object probability map (MNP) [65], histogram-based contrast (HC) [2], region covariance descriptors for visual saliency (COV) [67], spatially weighted dissimilarity (SWD) [64], looking beyond the image using its patches (LBI) [69], pattern distinctness using principal component analysis (PCA) [66], appearance similarity and spatial distribution (GC) [61], contextual hypergraph modelling (CHM) [62], absorbing Markov chain on an image graph model (MC) [63], region-based contrast (RC) [2].

5.2 Qualitative comparison

The results of qualitative comparison are displayed in [Fig. 9](#) and [Fig. 10](#). It is evident that the proposed method suppresses background regions and emphasizes foreground regions in an effective and uniform manner. In some cases, our approach and the RC [2] method show similar results. In most cases, however, our method is better in emphasizing the salient object due to pre-processing, global contrast enhancement, and histogram equalization. [Fig. 9](#) illustrates the comparisons between the first part of our method, which involves the saliency map based on global contrast, and five other methods using natural scene images with ground truth images that contain binary masks. These images contain typical natural scenes with various numbers of objects. It is also worth pointing out that our binary mask performs on par with the ground truth. [Fig. 10](#) shows the comparisons of our saliency map, saliency cuts result, and five other methods with the ground truth.



(a) Image (b) CA (c) GC (d) CHM (e) MC (f) HC (g) RC (h) Ours (i) GT
Fig. 9. Visual comparison between conventional approaches, the proposed method, and the ground truth (GT): (a) input image, (b) context-aware saliency (CA) [60], (c) appearance similarity and spatial distribution (GC) [61], (d) contextual hypergraph modelling (CHM) [62], (e) absorbing Markov chain on an image graph model (MC) [63], (f) histogram-based contrast (HC) [2], (g) region-based contrast (RC) [2], (h) our method, (i) ground truth (GT).



(a) Image (b) SWD (c) COV (d) SS (e) MNP (f) LBI (g) PCA (h) Ours (i) Ours (j) GT

Fig. 10. Visual comparison between other methods, our saliency map, and the binary mask: (a) input image, (b) spatially weighted dissimilarity (SWD) [64], (c) region covariance descriptors for visual saliency (COV) [67], (d) highlighting sparse saliency regions (SS) [68], (e) patch distinctness with an object probability map (MNP) [65], (f) looking beyond the image using its patches (LBI) [69], (g) pattern distinctness using principal component analysis (PCA) [66], (h) proposed global contrast based saliency map, (i) proposed saliency cuts, (j) ground truth (GT).

In addition, we performed experiments to evaluate our method in terms of its ability to generate salient object contours and inner edges. **Fig. 11** compares the experimental results of the proposed approach using various natural scene images. The proposed approach obtains a full resolution binary mask of high quality and accuracy, which is almost identical to the ground truth in most cases. However, with an excessive amount of contrast, the proposed method may fail to detect a salient object that has a color intensity similar to that of the background.

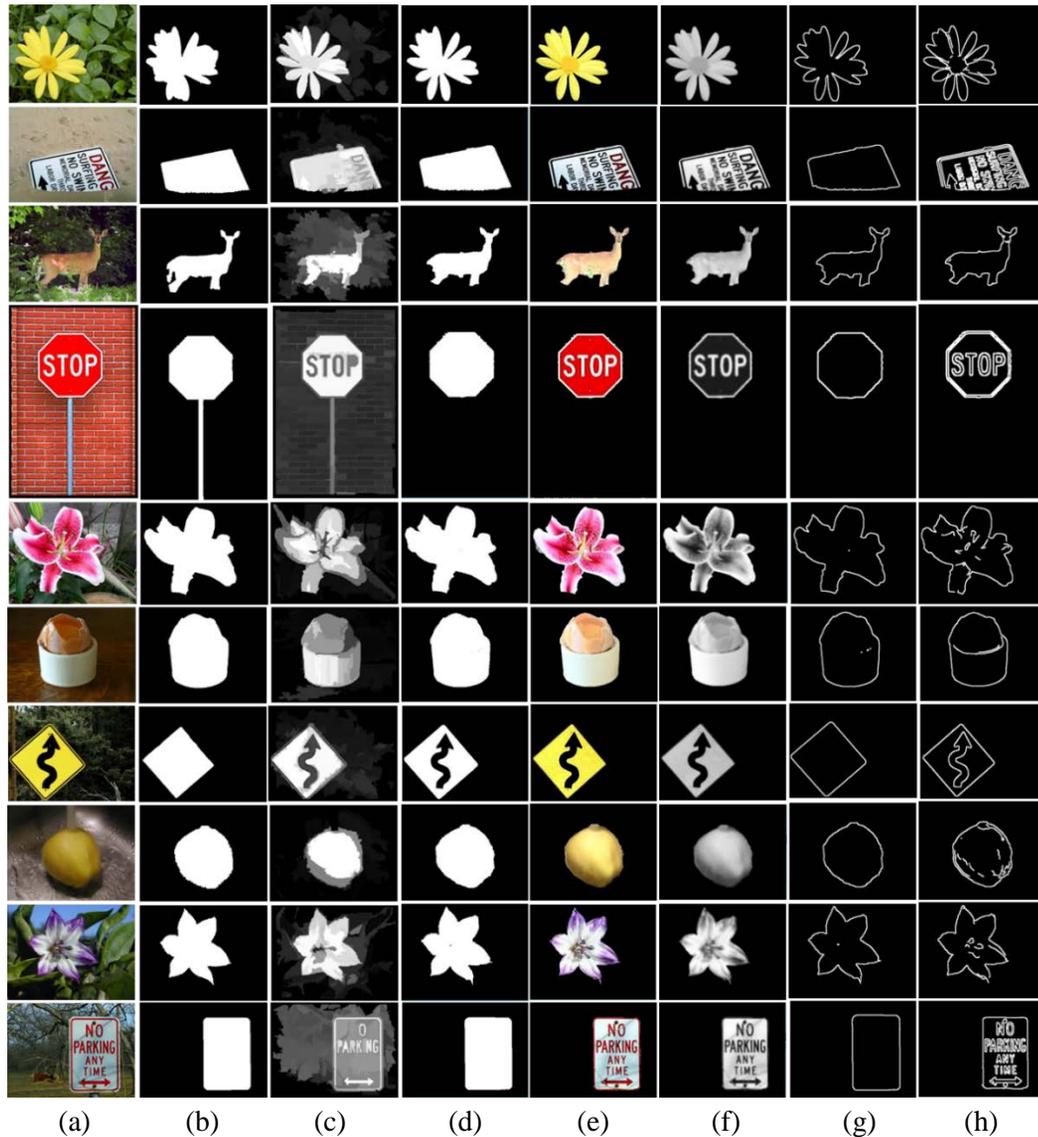


Fig. 11. Experimental results of the proposed method: (a) input image, (b) ground truth, (c) our saliency map, (d) our binary mask using saliency cuts, (e) full color space salient object, (f) bilateral filtering, (g) outer contour, (h) outer contour and inner edges.

Table 1 compares the processing time of the proposed method with those of other methods using images in an MSRA10K dataset (typical image resolution: 400×300). As shown in the table, CA, CHM, MNP, COV, SS, LBI are Matlab implementations, while GC, MC, SWD, PCA, HC, RC and our method are implemented using C++. The proposed method is

Table 1. Comparison of processing time required for each method.

Method	CA[53]	GC[54]	CHM[56]	MC[69]	SWD[58]	MNP[60]	COV[66]
Time (s)	45.3	0.056	17.4	0.235	0.163	21.6	26.2
Code Type	Matlab	C++	Matlab	C++	C++	Matlab	Matlab
Method	PCA[65]	LBI[68]	HC[2]	RC[2]	SS[67]	Ours	
Time (s)	4.74	252.4	0.019	0.154	0.83	0.157	
Code Type	C++	Matlab	C++	C++	Matlab	C++	

slower than GC, HC, and RC because it requires global contrast enhancement and image segmentation. However, the proposed method produces a high-quality, full-resolution saliency map.

6. Conclusion and Future work

This paper proposed a saliency region extraction technique based on global color contrast enhancement using histogram equalization and saliency region extraction, as well as saliency cuts based on adaptive triple thresholding. The proposed method highlights a saliency object and suppresses the background. We evaluated true color space saliency object extraction using a high-quality binary mask. With the ultimate objective of obtaining saliency object's contour and inner edges, we applied masking, bilateral filtering, and canny edge detection. We performed experiments to evaluate our approach based on a widely used benchmark database and achieved better results than other methods. Moreover, we were able to translate visually saliency objects into assistive technology systems for the blind, enabling them to easily recognize natural scene images. Because it can be difficult for the visually impaired to recognize an object with complex boundaries and unnecessary edges by touching the outer boundary, we intend to further develop efficient methods for simplifying inner edges and adding zooming features to facilitate edge detection. Moreover, the proposed approach can be used to extract pedestrians or moving objects from image information for autonomous driving.

Acknowledgment

This research was made possible by a grant from the R&D Program administered by the Korea Railroad Research Institute.

References

- [1] Z. Liu, O. Le Meur and Sh. Luo, "Superpixel-based saliency detection," in *Proc. of 14th Int. Workshop on Image and Audio Analysis for Multimedia Interactive Services*, pp. 1-4, July 3-5, 2013. [Article\(CrossRef Link\)](#)
- [2] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang and Sh. M. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569-582, March, 2015. [Article\(CrossRef Link\)](#)
- [3] Y. F. Ma and H. J. Zhang, "Contrast-based Image Attention Analysis by Using Fuzzy Growing," in *Proc. of the 11th ACM Int. Conf. on Multimedia*, pp. 374-381, November 2-8, 2003. [Article\(CrossRef Link\)](#)
- [4] T. Zhao, L. Li, X. Ding, Y. Huang and D. Zeng, "Saliency Detection with Spaces of Background-Based Distribution," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 683-687, May, 2016. [Article\(CrossRef Link\)](#)
- [5] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 185-207, January, 2012. [Article\(CrossRef Link\)](#)
- [6] M. Donoser, M. Urschler, M. Hirzer and H. Bischof, "Saliency Driven Total Variation Segmentation," in *Proc. of IEEE 12th Int. Conf. on Computer Vision*, pp. 814-824, September 29-October 2, 2009. [Article\(CrossRef Link\)](#)
- [7] G. Li, Y. Xie, L. Lin and Y. Yu, "Instance-Level salient object segmentation," *arXiv preprint*, April 12, 2017. [Article\(CrossRef Link\)](#)
- [8] Y. Gao, M. Shi, D. Tao and Ch. Xu, "Database Saliency for Fast Image Retrieval," *IEEE Transactions on Multimedia*, vol. 17, no. 3, pp. 359-369, March, 2015. [Article\(CrossRef Link\)](#)
- [9] X. Wei, Z. Tao, Ch. Zhang and X. Cao, "Structured Saliency Fusion Based on Dempster-Shafer Theory," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1345-1349, September, 2015. [Article\(CrossRef Link\)](#)
- [10] M. Wang, R. Hong, X. T. Yuan, Sh. Yan and T. S. Chua, "Movie2Comics: Towards a Lively Video Content Presentation," *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 858-870, June, 2012. [Article\(CrossRef Link\)](#)
- [11] M. Rabbani and R. Joshi, "An overview of the JPEG2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, no. 1, pp. 3-48, January, 2002. [Article\(CrossRef Link\)](#)
- [12] R. Gallea, E. Ardizzone and R. Pirrone, "Physical Metaphor for Streaming Media Retargeting," *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 971-979, June, 2014. [Article\(CrossRef Link\)](#)
- [13] U. Rutishauser, D. Walther, Ch. Koch and P. Perona, "Is bottom-up attention useful for object recognition?," in *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. II-37-II-44, June 27-July 2, 2004. [Article\(CrossRef Link\)](#)
- [14] J. Lei, C. Zhang, Y. Fang, Z. Gu, N. Ling and Ch. Hou, "Depth Sensation Enhancement for Multiple Virtual View Rendering," *IEEE Transactions on Multimedia*, vol. 17, no. 4, pp. 457-469, April, 2015. [Article\(CrossRef Link\)](#)
- [15] G. X. Zhang, M. M. Cheng, Sh. M. Hu and R. Martin, "A Shape-Preserving Approach to Image Resizing," *Journal compilation: The Eurographics Association and Blackwell Publishing Ltd*, vol. 28, no. 7, pp. 1897-1906, October, 2009. [Article\(CrossRef Link\)](#)

- [16] M. M. Cheng, F. L. Zhang, N. J. Mitra, X. Huang and Sh.-M. Hu, "RepFinder: Finding Approximately Repeated Scene Elements for Image Editing," *ACM Transactions on Graphics*, vol. 29, no. 4, pp. 83, July, 2010. [Article\(CrossRef Link\)](#)
- [17] R. Hong, M. Wang, G. Li and X. T. Yuan, "iComics. Automatic conversion of movie into comics," in *Proc. of 18th ACM Int. Conf. on Multimedia*, pp. 1599-1602, October 25-29, 2010. [Article\(CrossRef Link\)](#)
- [18] R. Hong, J. He, H. Zhang and T. S. Chua, "Mental visual indexing: Towards Fast Video Browsing," in *Proc. of the 2016 ACM on Multimedia Conf.*, pp. 621-625, October 15-19, 2016. [Article\(CrossRef Link\)](#)
- [19] L. Nie, R. Hong, L. Zhang, Y. Xia, D. Tao and N. Sebe, "Perceptual attributes optimization for multivideo summarization," *IEEE Transactions on Cybernetics*, vol. 46, no. 12, pp. 2991-3003, December, 2016. [Article\(CrossRef Link\)](#)
- [20] Ch. Cheng, Sh. Li, Y. Wang, H. Qin and A. Hao, "Video saliency detection via spatial-temporal fusion and low-rank coherency diffusion," *IEEE Transactions on Image Processing*, February, 2017. [Article\(CrossRef Link\)](#)
- [21] Q. Yan, L. Xu, J. Shi and J. Jia, "Hierarchical Saliency Detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1155-1162, June 23-28, 2013. [Article\(CrossRef Link\)](#)
- [22] J. Jungil, Y. Hongchan, L. Hyelim and C. Jinsoo, "Graphic haptic electronic board-based education assistive technology system for blind people," in *Proc. of IEEE Int. Conf. on Consumer Electronics*, pp. 364-365, January 9-12, 2015. [Article\(CrossRef Link\)](#)
- [23] R. S. Srivatsa and R. V. Babu, "Saliency object detection via objectness measure," in *Proc. of IEEE Int. Conf. on Image Processing*, pp. 4481-4485, September 27-30, 2015. [Article\(CrossRef Link\)](#)
- [24] J. Shi, Q. Yan, L. Xu and J. Jia, "Hierarchical Image Saliency Detection on Extended CSSD," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 717-729, April, 2016. [Article\(CrossRef Link\)](#)
- [25] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Human Neurobiol.*, vol. 4, no. 4, pp. 219-227, 1985. [Article\(CrossRef Link\)](#)
- [26] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259, November, 1998. [Article\(CrossRef Link\)](#)
- [27] J. Harel, C. Koch and P. Perona, "Graph-based visual saliency," in *Proc. of 19th Int. Conf. on Neural Information Processing Systems*, pp. 545-552, December 4-7, 2006. [Article\(CrossRef Link\)](#)
- [28] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, "Frequency-tuned saliency region detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1597-1604, June 20-25, 2009. [Article\(CrossRef Link\)](#)
- [29] M. Du, X. Wu, W. Chen and J. Wang, "Fusing Region Contrast and Graph Regularization for Saliency Detection," in *Proc. of 2015 27th Chinese Control and Decision Conf.*, pp. 5789-5794, May 23-25, 2015. [Article\(CrossRef Link\)](#)
- [30] T. Liu, J. Sun, N. N. Zheng, X. Tang and H. Y. Shum, "Learning to Detect A Saliency Object," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 353-367, February, 2011. [Article\(CrossRef Link\)](#)
- [31] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *Int. Journal of Computer Vision*, vol. 59, no. 2, pp. 167-181, September, 2004. [Article\(CrossRef Link\)](#)

- [32] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, May, 2002. [Article\(CrossRef Link\)](#)
- [33] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Susstrunk, "SLIC Superpixels," *EPFL*, no. Technical Report 149300, June, 2010. [Article\(CrossRef Link\)](#)
- [34] J. Lou, M. Ren and H. Wang, "Regional principal color based saliency detection," *PLoS ONE*, vol. 9, no. 11, pp. 1-13, November, 2014. [Article\(CrossRef Link\)](#)
- [35] F. Perazzi, P. Krahenbuhl, Y. Pritch and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 733-740, June 16-21, 2012. [Article\(CrossRef Link\)](#)
- [36] M. Hua, X. Bie and W. Wang, "Contrast based saliency detection using representative background priors," in *Proc. IEEE Int. Conf. on Image Processing*, pp. 3422-3425, September 15-18, 2013. [Article\(CrossRef Link\)](#)
- [37] J. Han, D. Zhang, X. Hu, L. Guo, J. Ren and F. Wu, "Background Prior-Based Salient Object Detection via Deep Reconstruction Residual," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 8, August, 2015. [Article\(CrossRef Link\)](#)
- [38] Q. Jiang, F. Shao and G. Jiang, "Efficient saliency detection using color contrast and similarity distribution information," in *Proc. of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, December 9-12, 2014. [Article\(CrossRef Link\)](#)
- [39] Q. Zhou, J. Chen, Sh. Ren, Y. Zhou, J. Chen and W. Liu, "On Contrast Combinations for Visual Saliency Detection," in *Proc. of IEEE Int. Conf. on Image Processing*, pp. 2665-2669, September 15-18, 2013. [Article\(CrossRef Link\)](#)
- [40] J. Zhou, J. Zhai and Y. Ren, "Salient Object Detection via Adaptive Region Merging," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 9, pp. 4386-4404, September, 2016. [Article\(CrossRef Link\)](#)
- [41] X. Qian, J. Han, G. Cheng and L. Guo, "Optimal contrast based saliency detection," *Pattern Recognition Letters*, vol. 34, no. 11, pp. 1270-1278, August 1, 2013. [Article\(CrossRef Link\)](#)
- [42] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1-8, June 17-22, 2007. [Article\(CrossRef Link\)](#)
- [43] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 185-198, January, 2010. [Article\(CrossRef Link\)](#)
- [44] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," *Association for Computing Machinery Transactions on Graphics*, vol. 23, no. 3, pp. 309-314, August, 2004. [Article\(CrossRef Link\)](#)
- [45] S. Li, R. Ju, T. Ren and G. Wu, "Saliency cuts based on adaptive triple thresholding," in *Proc. of IEEE Int. Conf. on Image Processing*, pp. 4609-4613, September 27-30, 2015. [Article\(CrossRef Link\)](#)
- [46] P. Mehrani and O. Veksler, "Saliency Segmentation based on Learning and Graph Cut Refinement," in *Proc. of British Machine Vision Conf.*, pp. 1-12. August 31-September 3, 2010. [Article\(CrossRef Link\)](#)
- [47] Y. Fu, J. Cheng, Z. Li, and H. Lu, "Saliency Cuts: An automatic approach to object segmentation," in *Proc. of 19th Int. Conf. Pattern Recognition*, pp. 1-4, December 8-11, 2008. [Article\(CrossRef Link\)](#)

- [48] Ç. Aytekin, E. C. Ozan, S. Kiranyaz and M. Gabbouj, "Visual saliency by extended quantum cuts," in *Proc. of Int. Conf. on Image Processing*, pp. 1692-1696, September 27-30, 2015. [Article\(CrossRef Link\)](#)
- [49] J. Peng, J. Shen, Y. Jia and X. Li, "Saliency Cut in Stereo Images," in *Proc. of IEEE Int. Conf. on Computer Vision Workshops*, pp. 22-28, December 2-8, 2013. [Article\(CrossRef Link\)](#)
- [50] L. Chen, B. Guo and W. Sun, "Obstacle Detection System for Visually Impaired People Based on Stereo Vision," in *Proc. of IEEE 4th Int. Conf. on Genetic and Evolutionary Computing*, pp. 723-726, December 13-15, 2010. [Article\(CrossRef Link\)](#)
- [51] L. Muthulakshmi and A. B. Ganesh, "Bimodal Based Environmental Awareness System for Visually Impaired People," in *Proc. of Int. Conf. on modeling, optimization and computing*, vol. 38, pp. 1132-1137, 2012. [Article\(CrossRef Link\)](#)
- [52] S. Wang, X. Yang and Y. Tian, "Detecting signage and doors for blind navigation and wayfinding," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 2, no. 2, pp. 81-93, July, 2013. [Article\(CrossRef Link\)](#)
- [53] N. Tagaki and J. Chen, "Character string extraction from scene images by eliminating non-character elements," in *Proc. of IEEE Int. Conf. on Systems*, pp. 3685-3690, October 5-8, 2014. [Article\(CrossRef Link\)](#)
- [54] N. Takagi, J. Chen, "A Broken Line Classification Method of Mathematical Graphs for Automating Translation into Scalable Vector Graphic," in *Proc. of 43rd IEEE Int. Symposium on Multiple-Valued Logic*, pp. 71-76, May 22-24, 2013. [Article\(CrossRef Link\)](#)
- [55] J. Chen and N. Takagi, "A Pattern Recognition Method for Automating Tactile Graphics Translation from Hand-Drawn Maps," in *Proc. of IEEE Int. Conf. on Systems*, pp. 4173-4178, October 13-16, 2013. [Article\(CrossRef Link\)](#)
- [56] S. S. Pathak, P. Dahiwale and G. Padole, "A combined effect of local and global method for contrast image enhancement," in *Proc. of IEEE Int. Conf. on Engineering and Technology*, pp. 1-5, March 20-25, 2015. [Article\(CrossRef Link\)](#)
- [57] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, January, 1979. [Article\(CrossRef Link\)](#)
- [58] S. Paris, P. Kornprobst, J. Tumblin and F. Durand, "Bilateral Filtering: Theory and Applications," *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1-73, August, 2009. [Article\(CrossRef Link\)](#)
- [59] G. Jie and L. Ning, "An improved adaptive threshold canny edge detection algorithm," in *Proc. of Int. Conf. on Computer Science and Electronics Engineerin*, pp. 164-168, March 23-25, 2012. [Article\(CrossRef Link\)](#)
- [60] S. Goferman, L. Z. Manor and A. Tal, "Context-aware saliency detection," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 34, no. 10, pp. 1915-1926, October, 2012. [Article\(CrossRef Link\)](#)
- [61] M. M. Cheng, J. Warrell, W. Y. Lin, S. Zheng, V. Vineet and N. Crook, "Efficient salient region detection with soft image abstraction" in *Proc. of IEEE Int. Conf. on Computer Vision*, pp. 1529-1536, December 1-8, 2013. [Article\(CrossRef Link\)](#)
- [62] X. Li, Y. Li, C. Shen, A. R. Dick and A. V. D. Hengel, "Contextual hypergraph modeling for salient object detection" in *Proc. of IEEE Int. Conf. on Computer Vision*, pp. 3328-3335, December 1-8, 2013. [Article\(CrossRef Link\)](#)
- [63] B. Jiang, L. Zhang, H. Lu, C. Yang and M. H. Yang, "Saliency detection via absorbing Markov chain," in *Proc. of IEEE Int. Conf. on Computer Vision*, pp. 1665-1672, December 1-8, 2013. [Article\(CrossRef Link\)](#)

- [64] L. Duan, C. Wu, J. Miao, L. Qing and Y. Fu, "Visual saliency detection by spatially weighted dissimilarity," in *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 473–480, June 20-25, 2011. [Article\(CrossRef Link\)](#)
- [65] R. Margolin, L. Z. Manor and A. Tal, "Saliency for image manipulation," *The Visual Computer*, vol. 29. no. 5, pp. 381-392, May 2013. [Article\(CrossRef Link\)](#)
- [66] R. Margolin, A. Tal and L. Z. Manor, "What makes a patch distinct?," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1139–1146, June 23-28, 2013. [Article\(CrossRef Link\)](#)
- [67] E. Erdem and A. Erdem, "Visual saliency estimation by nonlinearly integrating features using region covariances," *Journal Vision*, vol. 13, no. 4, pp. 1-20, March, 2013. [Article\(CrossRef Link\)](#)
- [68] X. Hou, J. Harel and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 1, pp. 194-201, January, 2012. [Article\(CrossRef Link\)](#)
- [69] P. Siva, C. Russell, T. Xiang and L. Agapito, "Looking beyond the image: Unsupervised learning for object saliency and detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 3238–3245, June 23-28, 2013. [Article\(CrossRef Link\)](#)



Hongchan Yoon received his B.S. degree in Computer Engineering from Gachon University in 2014, and M.S. degree in IT Convergence Engineering from Gachon University in 2016. He currently heads a team at the Software Development Division of PCT Inc. in Seongnam, Korea. His current research interests include image processing, object extraction, pattern recognition, and embedded systems.



Baek-Hyun Kim received his B.S., M.S. and Ph.D. degree from Inha University, Inchon, Korea in 1994, 1996 and 2003, under the Korean Air Line Scholarship Grants, respectively. Since 2003 he has been with Korea railroad Research Institute, as a senior researcher. His research interests include signal/image processing, multiple access communication systems, railway communication, and PRT(Personal Rapid Transit) network operation.



Mukhiddinov Mukhriddin received his B.S. in Informatics and Information Technologies from Tashkent University of Information Technologies in Tashkent, Uzbekistan in June 2015. He is currently pursuing a master's degree at the Department of IT Convergence Engineering in Gachon University in Seongnam, Korea. His current research interests include object extraction, image processing, pattern recognition, and embedded systems.



Jinsoo Cho received his B.S. in Electronic Engineering from Inha University in 1994, M.S. in Electrical Engineering from Columbia University in 1998, and Ph.D. in Electrical and Computer Engineering from the Georgia Institute of Technology in 2003. From 2004 to 2006, he was a senior research engineer in the D-TV development team of System LSI Division, Samsung Electronics Co., Ltd in Korea. He is currently an associate professor in the Department of Computer Engineering, College of IT, Gachon University in Seongnam, Korea. His current research interests include image/video processing, computer vision, and assistive technology for the visually impaired.