

PAPER

Co-Propagation with Distributed Seeds for Salient Object Detection

Yo UMEKI^{†a)}, Taichi YOSHIDA^{††b)}, *Nonmembers*, and Masahiro IWAHASHI^{†c)}, *Member*

SUMMARY In this paper, we propose a method of salient object detection based on distributed seeds and a co-propagation of seed information. Salient object detection is a technique which estimates important objects for human by calculating saliency values of pixels. Previous salient object detection methods often produce incorrect saliency values near salient objects in the case of images which have some objects, called the leakage of saliencies. Therefore, a method based on a co-propagation, the scale invariant feature transform, the high dimensional color transform, and machine learning is proposed to reduce the leakage. Firstly, the proposed method estimates regions clearly located in salient objects and the background, which are called as seeds and resultant seeds, are distributed over images. Next, the saliency information of seeds is simultaneously propagated, which is then referred as a co-propagation. The proposed method can reduce the leakage caused because of the above methods when the co-propagation of each information collide with each other near the boundary. Experiments show that the proposed method significantly outperforms the state-of-the-art methods in mean absolute error and F-measure, which perceptually reduces the leakage.

key words: *salient object detection, label propagation, machine learning, SIFT, HDCT*

1. Introduction

Saliency detection is a technique which detects salient local regions in images and widely used as pre-processing in image processing [1]–[7]. Salient regions are defined as attractive areas which have characteristics for human eyes such as high contrast, unique orient, distinctive color, and so on. Hence, saliency detection is used in estimating of human eye fixation and content-aware image compression [8]. Recently, several methods of saliency detection have been proposed based on machine learning and accurately detect attention regions for human [5], [7].

As derivation of saliency detection, salient object detection has been recently studied and its methods are efficient for image retargeting, recognition, and so on [9]–[24]. Salient object detection is a technique which is not only detecting salient regions but also objects such as a tall man, a

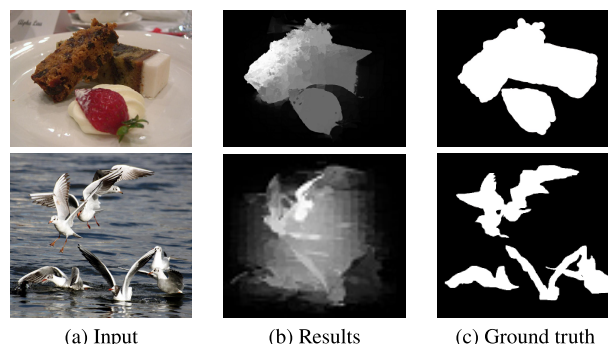


Fig. 1 Results of [16].

red car, a sign, and others. Since some image applications require the detection of salient objects [25]–[30], saliency detection is unsuitable for this requirement and salient object detection has been proposed on behalf of saliency detection. Thus, salient object detection has been proposed to recognize the salient objects, where both of them have different purpose and measures, which makes them difficult to fairly compared each other [31].

Some methods of salient object detection are realized on two-step methods and yields better results than ones of other methods. The methods are usually based on super-pixel segmentation. At first, a method of super-pixel segmentation is applied to input images, where it divides images into local clusters of pixels, called super-pixels [32]–[34]. The super-pixels which belong to the salient and background regions are detected, and called as salient and background seeds in this paper. Next, labels of seeds are propagated throughout images based on the features of super-pixels.

Although the two-step methods above can efficiently detect salient objects, they often produce incorrect labels near the boundary of objects especially in complex images. In this paper, the complex images are defined as some salient objects exist in one image, whereas simple images are defined as images which contain one big object per image. As the features of the object and the background are similar when it is near the boundary of the image, a label is leaked over the boundary, as shown in the bottom part of Fig. 1 (b). The phenomenon is referred as the label leakage and the limitations of the phenomenon is overcome in this paper to get a more reliable detection method.

We believe that in order to reduce the leakage in the label, it is effective to simultaneously propagated the salient

Manuscript received December 22, 2017.

Manuscript revised February 23, 2018.

Manuscript publicized March 9, 2018.

[†]The authors are with the Dept. of Electrical Electronics and Information Engineering, Nagaoka Univ. of Tech., Nagaoka-shi, 940–2137 Japan.

^{††}The author is with the Dept. of Communication Engineering and Informatics, Univ. of Electro-Commun, Chofu-shi, 182–8585 Japan.

a) E-mail: umeki@stn.nagaokaut.ac.jp

b) E-mail: t-yoshida@uec.ac.jp

c) E-mail: iwahashi@vos.nagaokaut.ac.jp

DOI: 10.1587/transinf.2017EDP7413

and background labels and distributed the seeds over the image. This is because the similarity between two super-pixels across the object boundary where the label leakage occurs is generally low. Then, if super-pixels across the object boundary have different labels, their influence through the propagation has reduced each other on the boundary. The above situation can be realized if the seeds are accurately distributed near the boundary. The label leakage is found to be reduced by using this method.

Hence, we proposed a co-propagation method with distributed seeds to realize the above methods in this paper. The proposed method sets label of salient and background seeds to positive and negative, respectively, and iteratively performs the propagation [16], [35], [36], which is called the co-propagation. Salient and background seeds are detected based on high dimensional color transform (HDCT) and scale invariant feature transform (SIFT). Moreover, the reliable labels, super-pixels are registered as seeds through the iterative propagation. Therefore, our proposal based on the above methods reduces the label leakage.

The proposed method objectively and perceptually shows comparable and better results in the case of simple and complex images, respectively, and reduces the label leakage compared to the state-of-the-art methods as proved through the experiments. The Mean Absolute Error (MAE) and F-measure are used to compare those methods. In the case of simple images, the proposed method shows almost comparable results, objectively and perceptually. In contrast, the proposed method has better scores of measures and reduces the label leakage in the case of complex images. It is experimentally found that the efficacy of the proposed method for salient object detection is higher than other prevalent methods.

2. Related Works

Salient object detection are mainly divided into top-down and bottom-up algorithms. The top-down algorithm detects specific ordinary objects which are previously designated based on the machine learning technique. Other algorithms calculate saliency values of ordinary objects based on two-step methods. Methods of bottom-up algorithm usually outperform ones of top-down for salient object detection, and the proposed method belongs to the bottom-up algorithm.

The machine learning based methods of top-down algorithm have been proposed in [9], [12]–[14], [23], [24]. Kocak *et al.* proposed the super-pixel based method which is well-suited for salient objects via discriminative dictionaries [13]. Zhao *et al.* introduced the deep learning based method for low-contrast background images [14], meanwhile Yang *et al.* proposed the method with the conditional random field (CRF) and dictionaries [24].

The bottom-up algorithm is mainly realized on two-step methods [10], [11], [16], [17], [19], [21]. Propagate labels of seeds based on the iterative propagation with color similarity and the Markov absorption probability are done by [16], [17]. A HDCT with random forest based detec-

tion method accurately detects positions of salient objects for complex images [21]. Those are the current state-of-the-art methods in salient object detection, and show better results than other conventional methods.

3. Fundamental Method for Salient Object Detection

3.1 Machine Learning

Methods using the machine learning for salient object detection have been proposed [9], [18], [21], and we use a method [21] in this paper. They realize pre-learned models from various natural images with their saliency values as ground truth based on the machine learning algorithms, and detect salient objects by applying resultant models. The method calculates a feature vector of each super-pixel and feature vectors are input to the machine learning algorithm, where the feature vector consists of the HDCT, the histogram of oriented gradients (HOG), and singular values of the super-pixel [21], [37], [38]. As the learning algorithm, the method uses the random forest [39]. At the detection part, feature vectors of an input image are calculated and saliency values of super-pixels are estimated by applying the resultant model to its feature vector.

3.2 Feature Vector

Various feature vectors of super-pixels have been proposed to represent their characteristic for saliency detection. In this paper, we use a feature vector consists of indices, gradients, color values, color histogram, and contrast in the multi-color space [10], [21], [37], [40], called the HDCT feature vector. The elements of the HDCT feature vector are called as the location, color, texture, color histogram, and color contrast features.

The location feature consists of average values of normalized pixel coordinates along vertical and horizontal directions, while the colors feature consists of average values of pixels in the RGB, CIELab, and HSV color space. The texture feature consists of the pixel number, the HOG feature, and the singular value feature [37], [38], [40].

The histogram feature is measured by the chi-square distance of histogram values between current and other super-pixels in the RGB, CIELab, and HSV color space. The distance is formulated as

$$H_i = \sum_{j=1}^N \sum_{k=1}^B \frac{(h_{ik} - h_{jk})^2}{(h_{ik} + h_{jk})}, \quad (1)$$

where H_i is the RGB value of the histogram feature in the i -th super-pixel, N and B is the number of super-pixels and histogram bins, respectively, and h_{ik} is the RGB histogram value of k -th bin in the i -th super-pixel. Similarly, CIELab and HSV values of the histogram feature are calculated.

The contrast feature consists of the global contrast, the local contrast, and the element distribution in each component of the color feature. The global and local contrasts

are calculated based on the distance between current and other super-pixels, and the element distribution is realized by measuring the compactness of color values in terms of its spatial color variance [10]. Let G_i and L_i be the global and local contrast values of each color in the i -th super-pixel, respectively, and they are defined as follows:

$$\begin{aligned} G_i &= \sum_{j=1}^N (c_i - c_j)^2, \\ L_i &= \sum_{j=1}^N \omega_{ij} (c_i - c_j)^2, \\ \omega_{ij} &= \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_p^2} \|\mathbf{p}_i - \mathbf{p}_j\|_2^2\right), \end{aligned} \quad (2)$$

where c_i is a component of the color feature in the i -th super-pixel, Z_i is the normalization term, $\mathbf{p}_i \in [0, 1]^2$ denotes a vector of the location feature in the i -th super-pixel, σ_p is a parameter to control the influence of neighboring super-pixels.

3.3 Label Propagation

The propagation-based method iteratively diffuses labels based on similarities between neighboring super-pixels. Let $\mathbf{v}_t \in [0, 1]^N$ be a vector of label values in the t -th iteration ($t \in \mathbb{N}$), and it is defined as

$$\mathbf{v}_t = F_{\mathbf{G}}(\mathbf{A}\mathbf{v}_{t-1}), \quad (3)$$

where $\mathbf{v}_0$ is a starting state of $\mathbf{v}_t$ as

$$[\mathbf{v}_0]_i = \begin{cases} 1 & \text{if } i \in \mathbf{G} \\ 0 & \text{otherwise} \end{cases}, \quad (4)$$

$[\cdot]_i$ denotes the i -th element, $\mathbf{G}$ is a set of initial seed indices, $F_{\mathbf{G}}$ is a function which resets label values of seeds as

$$[F_{\mathbf{G}}(\mathbf{x})]_i = \begin{cases} 1 & \text{if } i \in \mathbf{G} \\ [\mathbf{x}]_i & \text{otherwise} \end{cases}, \quad (5)$$

and $\mathbf{A}$ is an affinity matrix. $\mathbf{A}$ is defined as

$$\mathbf{A} = \mathbf{D}^{-1}\mathbf{W}, \quad (6)$$

where $\mathbf{D} = \text{diag}(q_1, q_2, \dots, q_N)$, diag means a diagonal matrix, $q_i = \sum_{j=1}^N s_{ij}$, $\mathbf{W}$ denotes a $N \times N$ matrix whose (i, j) element is s_{ij} , and s_{ij} is a similarity value between i -th and j -th super-pixels. s_{ij} is defined as

$$s_{ij} = \begin{cases} \exp\left(-\frac{d(\mathbf{f}_i, \mathbf{f}_j)}{\sigma^2}\right) & \text{if } j \in \mathcal{N}(i) \text{ or } i, j \in \mathcal{B} \\ 0 & \text{if } i = j \text{ or otherwise} \end{cases}, \quad (7)$$

where $\mathbf{f}_i$ denotes a vector which has color values of the i -th super-pixel in CIELab [41]. d is the Euclidean distance function, σ is a parameter to control the similarity, and $\mathcal{N}(i)$ and $\mathcal{B}$ represents the sets of super-pixel indices near the i -th super-pixel and in the image boundary. If the difference of $\mathbf{v}_t$ and $\mathbf{v}_{t-1}$ is smaller than the threshold value, the iteration

process is finished.

4. Proposed Method

4.1 Framework

In this paper, we propose the co-propagation method with distributed seeds for salient object detection. The proposed method reduces the label leakage in complex images. The co-propagation is when the labels simultaneously propagating the collision seeds. Distributed seeds in salient and background regions are detected by using feature vectors of the HDCT and densities of SIFT feature points, respectively.

The first step in the proposed method is detecting the seeds. Salient seeds are detected by using a model learned by the random forest algorithm with feature vectors of the HDCT [39]. The random forest algorithm is efficient in large databases and has a generalization ability. On the other hand, background seeds on regions whose densities of SIFT feature points are set low. As the regions which are blurred and have tedious colors are considered unimportant, thus we define them as a background in this proposal.

The second step of the proposed method is calculating the label values of super-pixels by the co-propagation. Salient and background labels are set to positive and negative on seeds, respectively. Label values are iteratively propagated according to color similarities. The iterative process is stopped when label values are slightly changed by one propagation. In the iteration process, if label values of super-pixels are relatively high, their indices are registered as seeds.

Figure 2 shows the overview of the proposed method. First, an input image is divided into super-pixels, then, salient and background seeds are detected. Both salient and background labels are then simultaneously propagated for all super-pixels by the co-propagation method. Final label values are outputted as resultant saliency values.

4.2 Estimation of Initial Seeds

As mentioned in the previous section, the initial salient seeds are detected by applying the pre-learned model for super-pixels. The model is realized by using random forest algorithm which is applied into the MSRA-B dataset that consists of lots of natural images. The HDCT feature is used as feature vector in this paper. The model is applied for super-pixels of an input image. We set salient seeds in super-pixels whose resultant rates are over θ_f , where θ_f is a parameter and is set as 0.75. $\mathbf{G}_f$ is a set of indices of salient seeds.

The initial seeds of background are detected based on the density of SIFT feature points. First, we find feature points in an input image by the SIFT key point detection. The feature point density of the i -th super-pixel g_i is calculated by dividing the number of feature points with the number of pixels in the i -th super-pixel. Let $\bar{g}$ be an average of g_i in all super-pixels. The background seeds in super-pixels

as $g_i \leq \bar{g} \times \theta_b$, where θ_b is a parameter and set 0.2 in this paper. Let $\mathbf{G}_b$ is a set of indices of background seeds.

4.3 Co-propagation of Labels

The co-propagation is a simultaneous propagation of labels. We introduce $F_{\mathbf{G}_f, \mathbf{G}_b}$ instead of $F_{\mathbf{G}}$ in (5) as

$$[F_{\mathbf{G}_f, \mathbf{G}_b}(\mathbf{x})]_i = \begin{cases} 1 & \text{if } i \in \mathbf{G}_f \\ -1 & \text{if } i \in \mathbf{G}_b \\ [\mathbf{x}]_i & \text{otherwise} \end{cases}. \quad (8)$$

The final saliency values are calculated by applying the iteration process in (3) and the result is normalized in $[0, 1]$.

4.4 Adding Seeds

In the proposed method, we add seeds at the regions which have high absolute label values in every n times of iteration. The proposed method updates $\mathbf{G}_f$ and $\mathbf{G}_b$ when $\text{mod}(t, n) = 0$ as

$$\begin{aligned} \mathbf{G}_f &= \{\mathbf{G}_f, i\} & \text{if } \tau_1 < [\mathbf{v}_t]_i \text{ and } \tau < s_{ij} \quad (j \in \mathbf{G}_f) \\ \mathbf{G}_b &= \{\mathbf{G}_b, i\} & \text{if } \tau_2 > [\mathbf{v}_t]_i \text{ and } \tau < s_{ij} \quad (j \in \mathbf{G}_b) \end{aligned}, \quad (9)$$

where τ , τ_1 , and τ_2 are threshold parameters. We set $\tau = \tau_1 = -\tau_2 = 0.7$, and s_{ij} is shown in (7). Figure 3 shows

the addition of background seeds into background regions. However, note that excessive seeds may cause false detection, thus, seeds are not added at every iteration.

5. Experiment

The proposed method is compared with the state-of-the-art methods for salient object detection. PASCAL-S,

Table 1 MAE scores of PASCAL-S [12].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Horse	0.046	0.121	0.030	0.026
Sheep	0.120	0.123	0.096	0.010
Cat	0.228	0.158	0.174	0.041
Cow	0.109	0.104	0.068	0.007
Dog	0.289	0.099	0.226	0.028
Animals	0.098	0.104	0.018	0.023
Average	0.166	0.135	0.158	0.116

Table 2 F-measure scores of PASCAL-S [12].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Horse	0.776	0.189	0.845	0.855
Sheep	0.585	0.573	0.683	0.840
Cat	0.496	0.681	0.648	0.922
Cow	0.639	0.661	0.789	0.902
Dog	0.556	0.871	0.684	0.936
Animals	0.469	0.427	0.896	0.864
Average	0.453	0.513	0.493	0.548

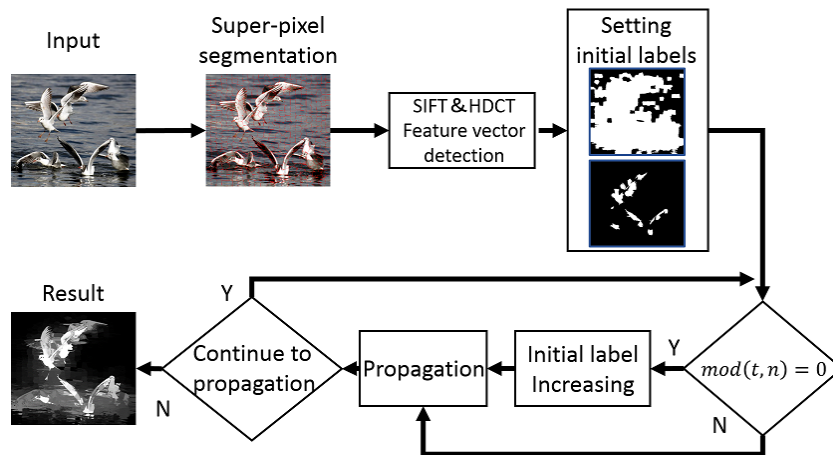


Fig. 2 Overview of proposed method.

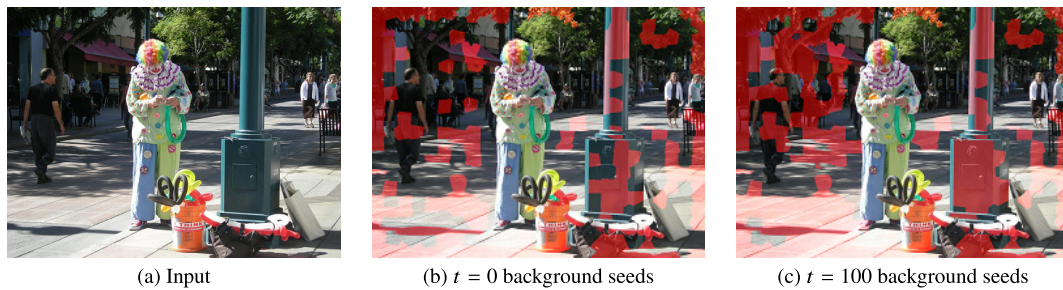


Fig. 3 Resultant seeds adding.

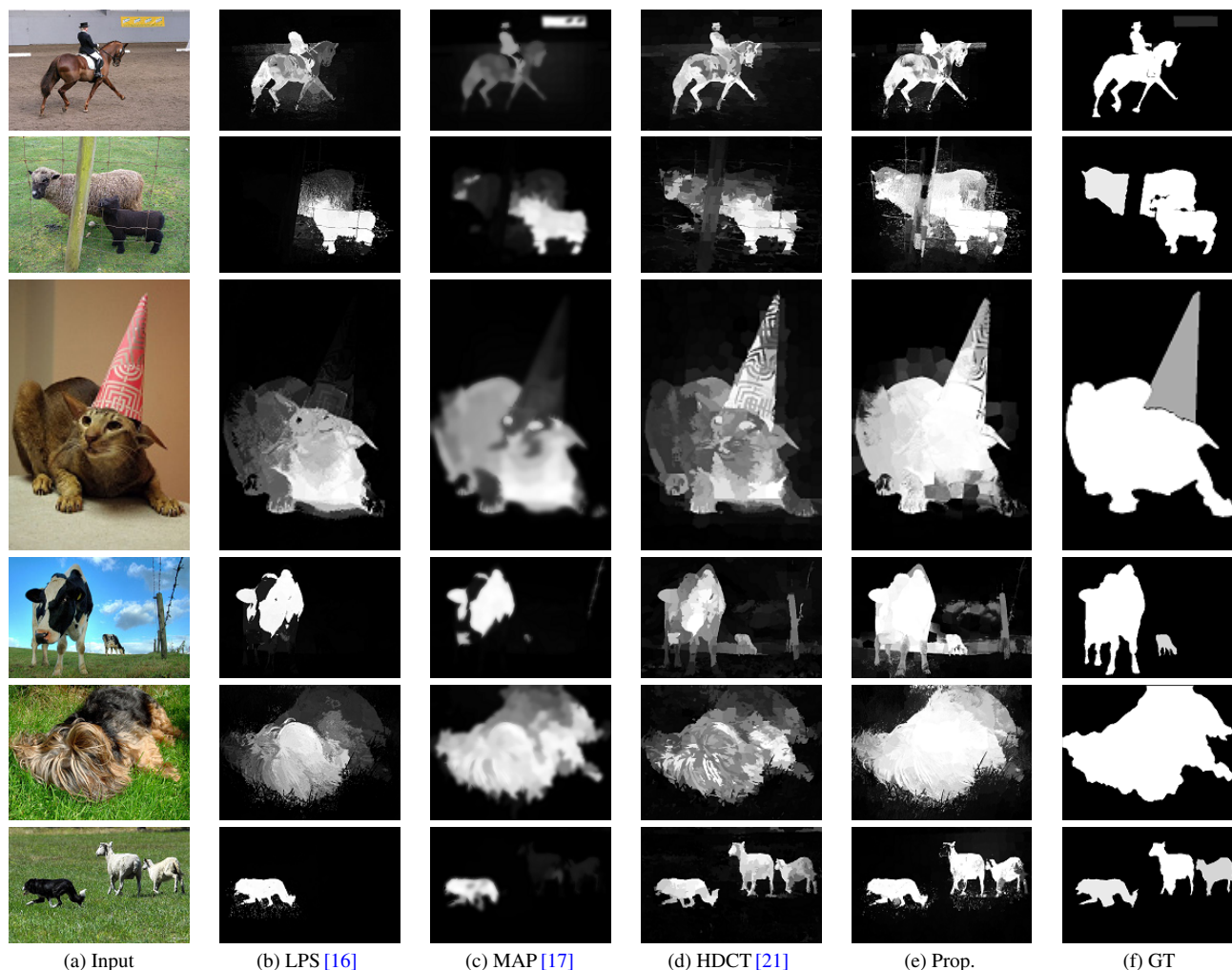


Fig. 4 Resultant saliency maps of PASCAL-S [12].

Table 3 MAE scores of MSRA10K [9].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Traffic light	0.152	0.155	0.096	0.078
Ice	0.108	0.058	0.012	0.010
Guitar	0.043	0.052	0.009	0.007
Bird	0.006	0.010	0.002	0.004
Average	0.085	0.069	0.084	0.066

Table 4 F-measure scores of MSRA10K [9].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Traffic light	0.278	0.331	0.659	0.729
Ice	0.464	0.768	0.959	0.961
Guitar	0.906	0.883	0.970	0.974
Bird	0.978	0.968	0.984	0.984
Average	0.708	0.750	0.736	0.756

MSRA10K and ECSSD datasets are used for test images [12], [42], [43]. Images of PASCAL-S dataset is categorized as complex ones with similar colors, many objects, and inconstant positions of object. MSRA10K and ECSSD datasets have simple images which only have one object on the center of image. The prevalent methods compared in this

Table 5 MAE scores of ECSSD [43].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Baseball	0.056	0.036	0.022	0.018
Woman	0.133	0.134	0.019	0.006
Garden	0.001	0.001	0.068	0.004
Squirrel	0.033	0.021	0.040	0.024
Average	0.129	0.092	0.140	0.100

Table 6 F-measure scores of ECSSD [43].

	LPS [16]	MAP [17]	HDCT [21]	Prop.
Baseball	0.792	0.681	0.902	0.909
Woman	0.623	0.619	0.955	0.981
Garden	0.989	0.992	0.735	0.981
Squirrel	0.915	0.945	0.902	0.909
Average	0.561	0.649	0.565	0.619

paper are the current state-of-the-art in two-step salient object detection, called MAP [17], LPS [16], and HDCT [21].

Since some methods are realized in super-pixel, we use the simple linear iterative clustering (SLIC) as a method of super-pixel segmentation and the number of super-pixels is set as 500 [33]. We set the parameter $n = 100$ at seeds

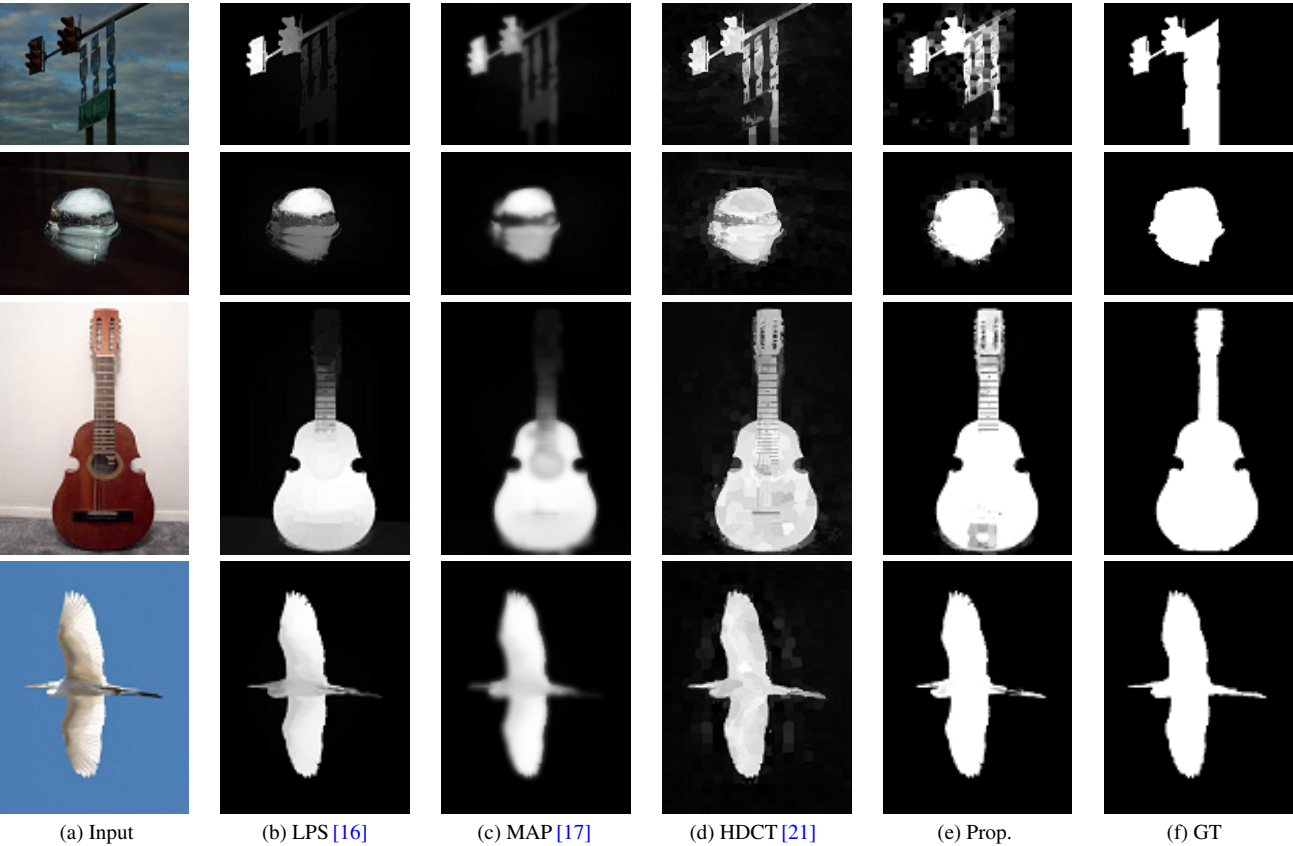


Fig. 5 Resultant saliency maps of MSRA10K [9].

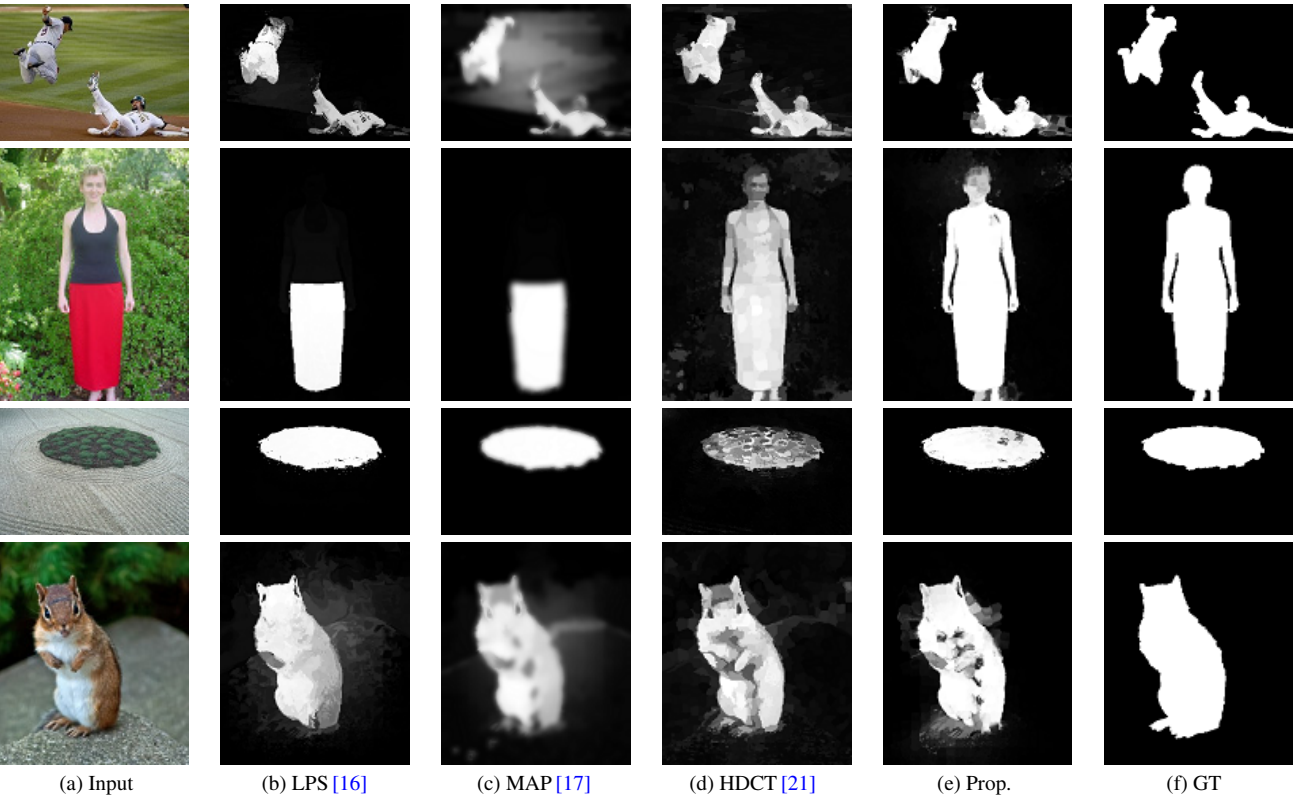


Fig. 6 Resultant saliency maps of ECSSD [43].

adding and the termination condition of iterative process at $\mathbf{v}_t - \mathbf{v}_{t-1} < 0.000001$.

The MAE and F-measure scores are shown in Table 1-6 where ‘Prop.’ refers to the proposed method, ‘Average’ refers to the average values of 850 images in Table 1 and 2, 10000 images in Table 3 and 4, 1000 images in Table 5 and 6. Table 1 and 2 show the results for complex images. It is proved that the proposed method almost outperforms other conventional methods in the implementation to the complex images. Meanwhile Table 3-6 shows the results for simple image datasets. The proposed method shows slightly better results compared with the conventional methods, and those in Table 3 and 4 and comparable and slightly worse ones from Table 5 and 6, respectively.

Some resultant saliency maps of PASCAL-S, MSRA10K, and ECSSD datasets are shown in Fig. 4, 5, and 6, where ‘GT’ means the ground truth. The proposed method shows uniformly high saliency values on objects and reduces the label leakage of positive and negative, compared with the LPS and MAP methods. MAP method particularly produces blurred edges of results and the leakage, which is unsuitable for salient object detection. The proposed method perceptually shows comparable results with HDCT for simple images, whereas for complex images, HDCT often shows low saliency values in objects such as ‘Sheep’, ‘Cat’, and ‘Dog’. Furthermore, the proposed method avoids losing objects for complex images such as ‘Sheep’ and ‘Animals’.

Based on these results, it is observed that the proposed method outperforms the state-of-the-art methods for salient object detection as mentioned in Sect. 1. The proposed method outperforms conventional methods for complex images, objectively and perceptually. The proposed method shows comparable and slightly better results, objectively and perceptually for simple images, but MAP method is better than the proposed method in certain images. However, MAP method is known to always produce blurred results, and is unsuitable for salient object detection. Therefore, the proposed method is found to be superior compared with state-of-the-art methods in salient object detection.

6. Conclusion

In this paper, we proposed a method of salient object detection based on the SIFT, HDCT, machine learning, and co-propagation. The proposed method estimated seeds of salient and background, which are based on machine learning with HDCT and the density of SIFT feature points. The co-propagation diffuses labels of seeds, and they are propagated to super-pixels according to the feature’s similarities. By using the two-step methods in salient object detection, our proposal produces better results than the state-of-the-art methods for complex images, and shows comparable results for simple images. The proposed method is also found to be effective for suppressing the leakage.

The future work of the proposed method is an accurate detection of background seeds. Figure 7 shows failure

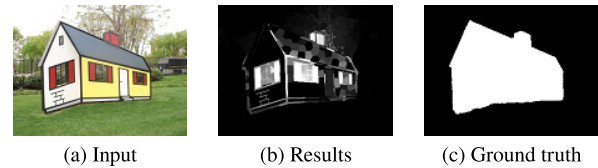


Fig. 7 Some failure cases of Prop.

cases of the proposed method. We defined that background regions are blurred and have tedious colors. Hence, the proposed method causes false detections of background seeds in objects which have features likely backgrounds. The proposed method cannot detect salient object using inaccurate seeds. Recently, convolutional neural networks based semantic segmentation method has been proposed [44]. They accurately detect regions such as sky, ground and so on. Those regions are mainly categorized in backgrounds and hence we can accurately detect background seeds. Detecting seeds accurately, the proposed method solves above problem.

References

- [1] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.20, no.11, pp.1254–1259, 1998.
- [2] L. Itti, “Automatic foveation for video compression using a neurobiological model of visual attention,” *IEEE Trans. Image Process.*, vol.13, no.10, pp.1304–1318, 2004.
- [3] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, “Frequency-tuned salient region detection,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.1597–1604, 2009.
- [4] M. Wang, J. Konrad, P. Ishwar, K. Jing, and H. Rowley, “Image saliency: From intrinsic to extrinsic context,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.417–424, 2011.
- [5] L. Wang, H. Lu, X. Ruan, and M.-H. Yang, “Deep networks for saliency detection via local estimation and global search,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.3183–3192, 2015.
- [6] J. Ren, X. Gong, L. Yu, W. Zhou, and M.Y. Yang, “Exploiting global priors for RGB-D saliency detection,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.25–32, 2015.
- [7] J. Zhang and S. Sclaroff, “Exploiting surroundedness for saliency detection: A boolean map approach,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.38, no.5, pp.889–902, 2016.
- [8] N. Bruce and J. Tsotsos, “Saliency based on information maximization,” *Proc. Advances Neural Information Process. Syst.*, pp.155–162, 2006.
- [9] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, “Learning to detect a salient object,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.33, no.2, pp.353–367, 2011.
- [10] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, “Saliency filters: Contrast based filtering for salient region detection,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.733–740, 2012.
- [11] X. Shen and Y. Wu, “A unified approach to salient object detection via low rank matrix recovery,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.1853–860, 2012.
- [12] Y. Li, X. Hou, C. Koch, J.M. Rehg, and A.L. Yuille, “The secrets of salient object segmentation,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.280–287, 2014.
- [13] A. Kocak, K. Cizmeciler, A. Erdem, and E. Erdem, “Top down saliency estimation via superpixel-based discriminative dictionaries,” *Proc. British Mach. Vis. Conf.*, 2014.
- [14] R. Zhao, W. Ouyang, H. Li, and X. Wang, “Saliency detection by multi-context deep learning,” *Proc. IEEE Conf. Comput. Vis. Patt.*

- Recognit., pp.1265–1274, 2015.
- [15] A. Borji, M.-M. Cheng, H. Jiang, and J. Li, “Salient object detection: A benchmark,” *IEEE Trans. Image Process.*, vol.24, no.12, pp.5706–5722, 2015.
 - [16] H. Li, H. Lu, Z. Lin, X. Shen, and B. Price, “Inner and inter label propagation: Salient object detection in the wild,” *IEEE Trans. Image Process.*, vol.24, no.10, pp.3176–3186, 2015.
 - [17] J. Sun, H. Lu, and X. Liu, “Saliency region detection based on markov absorption probabilities,” *IEEE Trans. Image Process.*, vol.24, no.5, pp.1639–1649, 2015.
 - [18] N. Tong, H. Lu, X. Ruan, and M.-H. Yang, “Salient object detection via bootstrap learning,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.1884–1892, 2015.
 - [19] M.-M. Cheng, N.J. Mitra, X. Huang, P.H.S. Torr, and S.-M. Hu, “Global contrast based salient region detection,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.37, no.3, pp.569–582, 2015.
 - [20] Y. Qin, H. Lu, Y. Xu, and H. Wang, “Saliency detection via cellular automata,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.110–119, 2015.
 - [21] J. Kim, D. Han, Y.-W. Tai, and J. Kim, “Salient region detection via high-dimensional color transform and local spatial support,” *IEEE Trans. Image Process.*, vol.25, no.1, pp.9–23, 2016.
 - [22] C. Sheth and R.V. Babu, “Object saliency using a background prior,” *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp.1931–1935, 2016.
 - [23] F. Yang and M.-H. Yang, “Top-down visual saliency via joint crf and dictionary learning,” *IEEE Trans. Patt. Anal. Mach. Int.*, pp.2296–2303, 2016.
 - [24] J. Yang and M.-H. Yang, “Top-down visual saliency via joint crf and dictionary learning,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.39, no.3, pp.576–588, 2017.
 - [25] U. Rutishauser, D. Walther, C. Koch, and P. Perona, “Is bottom-up attention useful for object recognition?,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, vol.2, pp.37–44, 2004.
 - [26] V. Setlur, S. Takagi, R. Rasker, M. Gleicher, and B. Gooch, “Automatic image retargeting,” *Proc. ACM Mobile ubiquitous multimedia*, pp.59–68, 2005.
 - [27] L. Marchesotti, C. Cifarelli, and G. Csurka, “A framework for visual saliency detection with applications to image thumbnailing,” *Proc. IEEE Int. Conf. Comput. Vis.*, pp.2232–2239, 2009.
 - [28] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, “A comparative study of image retargeting,” *ACM Trans. graphics*, vol.29, no.6, pp.160–9, 2010.
 - [29] X. Li, Y. Li, C. Shen, A. Dick, and A.V.D. Hengel, “Contextual hypergraph modeling for salient object detection,” *Proc. IEEE Int. Conf. Comput. Vis.*, pp.3328–3335, 2013.
 - [30] A. Jose and I. Heisterklaus, “Bag of fisher vectors representation of images by saliency-based spatial partitioning,” *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp.1762–1766, 2017.
 - [31] A. Kimura, R. Yonetani, and T. Hirayama, “Computational models of human visual attention and their implementations: A survey,” *IEICE Trans. Information and Syst.*, vol.96, no.3, pp.562–578, 2013.
 - [32] A. Levinstein, A. Stere, K.N. Kutulakos, D.J. Fleet, and S.J. Dickinson, “Turbopixels: Fast superpixels using geometric flows,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.31, no.12, pp.2290–2297, 2009.
 - [33] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “SLIC superpixels,” *Proc. EPFL Technical Report*, 149300, 15p, 2010.
 - [34] Z. Li and J. Chen, “Superpixel segmentation using linear spectral clustering,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.1356–1363, 2015.
 - [35] G.R. Shafer and P.P. Shenoy, “Probability propagation,” *Annals Math. Artificial Int.*, vol.2, no.1–4, pp.327–351, 1990.
 - [36] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” *Proc. Neural Information Process. Syst.*, pp.545–552, 2006.
 - [37] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, vol.1, pp.886–893, 2005.
 - [38] B. Su, S. Lu, and C.L. Tan, “Blurred image region detection and classification,” *Proc. ACM Int. Conf. Multimed.*, pp.1397–1400, 2011.
 - [39] L. Breiman, “Random forests,” *Mach. Learn.*, vol.45, no.1, pp.5–32, 2001.
 - [40] H. Andrews and C. Patterson, “Singular value decompositions and digital image processing,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol.24, no.1, pp.26–53, 1976.
 - [41] M.D. Fairchild, *Color Appearance Models*, Wiley, 2013.
 - [42] T. Judd, F. Durand, and A. Torralba, “A benchmark of computational models of saliency to predict human fixations,” *Proc. MIT Technical Reports*, MIT-CSAIL-TR-2012-001, 24p, 2012.
 - [43] J. Shi, Q. Yan, L. Xu, and J. Jia, “Hierarchical image saliency detection on extended cssd,” *IEEE Trans. Patt. Anal. Mach. Int.*, vol.38, no.4, pp.717–729, 2016.
 - [44] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, pp.3431–3440, 2015.



Yo Umeke received B.Eng. and M.Eng degrees from Nagaoka University of Technology, Nagaoka, Japan, in 2015 and 2017, respectively. He is currently a doctoral course student in the Department of Information Science and Control Engineering. His research interests are in the area of saliency detection.



Taichi Yoshida received B.Eng., M.Eng and Ph.D. degrees in Engineering from Keio University, Yokohama, Japan, in 2006, 2008 and 2013, respectively. From 2014 to 2017, he joined Nagaoka University of Technology. In 2018, he joined University of Electro-Communication where he is currently an Assistant Prof. in the Department of Electrical, Electronics and Information Engineering. His research interests are in the field of filter bank design and image coding application.



Masahiro Iwahashi received his B.E., M.E., and D.E. degrees in Electrical Engineering from Tokyo Metropolitan University in 1988, 1990, and 1996, respectively. In 1990, he joined Nippon Steel Co., Ltd. From 1991 to 1992, he was seconded to Graphics Communication Technology Co., Ltd. In 1993, he joined Nagaoka University of Technology, where he is currently a Professor in the Department of Electrical Engineering, Faculty of Technology. From 1995 to 2001, he was also a lecturer at the Nagaoka Technical College. From 1998 to 2001, he relocated to Thammasat University, Thailand, and to the Electronic Engineering Polytechnic Institute of Surabaya, Indonesia, as a JICA expert. His research interests are in the area of digital signal processing, multi-rate systems, and image compression. He served as an editorial committee member of the Transaction on Fundamentals of the IEICE from 2007 to 2011.