An Improved Hola Framework for Saliency Detection

1st Donyarut Kakanopas
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
k.donyarut@gmail.com

2nd Kuntpong Woraratpanya*
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
kuntpong@it.kmitl.ac.th

***Abstract*—Recently, a Hola framework was designed and implemented for saliency detection. It showed good performance among existing saliency detection methods. However, that proposed framework still has room for performance improvement. Hence, this paper proposed an improved Hola (i-Hola) framework for saliency detection. The new framework were improved in two important parts: (i) the Hola filter was redesigned for collaborating with a low-pass filter and (ii) the final saliency map selection was modified by using a combination of the first and second round saliency maps. The experimental results showed that the proposed framework can provide the improved performance of saliency detection.**

***Index Terms*—Saliency detection, improved Hola framework, Hola filter**

I. Introduction

Saliency detection aims to extract the area of interest of images. Recently, we have proposed a Hola filter in conjunction with salient detection framework, called a Hola framework for short, in detecting precisely uniform salient shape maps with sharp edge [1]. However, the Hola framework still has room for improvement in detecting saliency maps. Therefore, this paper focuses on the quality improvement of salient areas by modifying our previous Hola framework.

In our previous framework, as shown in Fig. 1(a) and reported in [1], six different Hola filters were designed based on amplitude spectrum filtering, responding to a low-frequency band (LFB) and median frequency band (MFB) to support a variety of image characteristics. The high brightness salient areas were extracted by a large LFB, while the low brightness salient areas were extracted by a small LFB. Also, the large MFB can extract salient textures. However, the appropriate LFB and MFB are challenging in obtaining uniformly salient regions with sharp edge. In addition, the previous Hola framework was defined the final saliency map by selecting from the minimum entropy of a set of enhanced candidate saliency maps in the second round of processing, thus causing the unsatisfied saliency areas in some cases as shown in Fig. 2.

In order to achieve the higher quality of saliency areas, an improved Hola (i-Hola) framework is proposed as shown in Fig. 1(b).

The rest of this paper is organized as follows. Section II points out the room for improvement of the previous Hola framework. The proposed method is described in section III. The experimental results and discussion are given in section IV. Finally, we conclude our work in section V.

*Corresponding author: kuntpong@it.kmitl.ac.th

II. Problem Statement

A previous Hola framework achieved with high quality of saliency maps, uniformly salient areas with sharp edge. However, some detected saliency maps of high contrast images with complex background, low contrast images, and bright background images still do not satisfy, since this framework cannot provide completely salient locations. In other words, some salient areas and details are lost. We tested three sample images, in different characteristics: a high contrast image, a low contrast image, and a bright background image, with the existing Hola framework. As a result, the second row of Fig. 2 shows the incorrect detected saliency map, whereas the first and third rows shows the unsatisfied saliency maps. The main causes of this problem were from (i) the inappropriate design of Hola filters and (ii) the failure of final saliency map selection.

III. Proposed Method

As mentioned in problem statement, we proposed an i-Hola framework for improving the quality of saliency maps. This framework was derived from the original one by (i) redesigning the Hola filter and (ii) modifying the final saliency map selection.

The proposed framework consists of tree main processes: (i) candidate saliency map extraction, (ii) candidate saliency map enhancement, and (iii) final saliency map selection as shown in Fig. 2(b).

A. Candidate Saliency Map Extraction

Based on an assumption that salient targets mostly hide in low frequency band; therefore, we improved the efficiency of candidate saliency map extraction by redesigning Hola filters in cooperation with low-pass filters to support low contrast and bright background images. Consequently, Fig. 1(b) shows the redesign of filters, consisting of two Hola filters and two low-pass filters. Inside processing of this block can be briefly described as follow: An RGB color image is converted to grayscale and LAB color space images. Each separate image is resized to 256×256 pixels, and then transformed to frequency (Fourier) domain. Next, i-Hola filters are used to suppress magnitude components. After that, the candidate saliency maps are obtained by taking an invert Fourier transform to the original phase and the suppressed magnitude components.

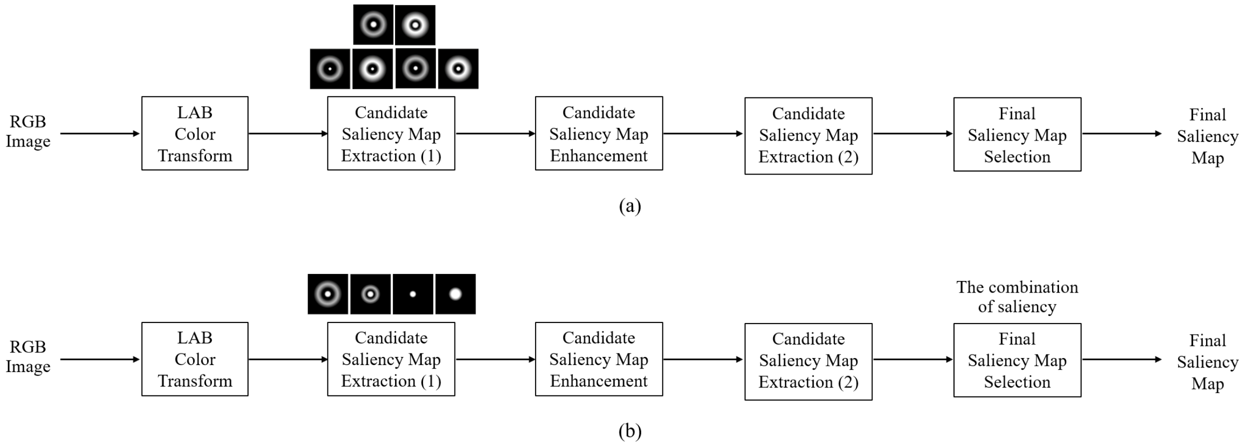

Fig. 1: Overall frameworks of (a) Hola and (b) i-Hola.

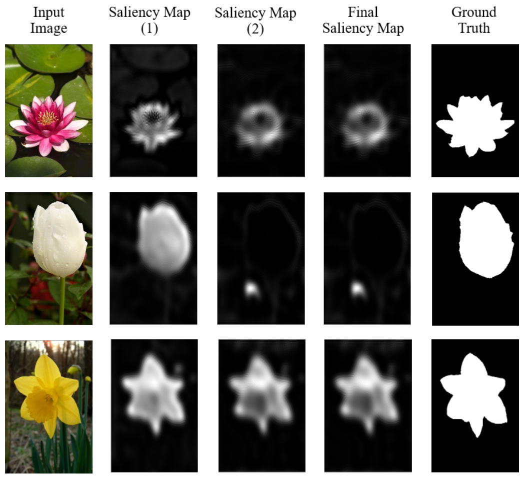

Fig. 2: An example of unsatisfied final saliency maps of our previous framework

B. Candidate Saliency Map Enhancement

In this process, candidate saliency maps of each separate image in the first round are enhanced by combining an edge image of gray channel from Prewitt algorithm. Each enhanced candidate saliency map is forwarded to candidate saliency map extraction process again, as the second round, to remove some non-saliency regions out. The obtained candidate saliency maps in the second round provide sharp edge more than their maps in the first round as shown in Fig. 2.

C. Final Saliency Map Selection

In order to select the optimal one in a set of all candidate saliency maps, the minimum entropy, as described in [1], is used to define the first and second round saliency maps. Then, the final saliency map, (S), is obtained from a combination of the first round saliency map, (s_1), and the second round saliency map, (s_2), as defined by Eq. (1) Lastly, the final saliency map image is scaled up to the original size.

$$S = (w_1 * s_1) + (w_2 * s_2) \qquad (1)$$

where w_1 and w_2 are the weight parameters. Here, w_1 and w_2 are set to 0.4 and 0.6, respectively. The output is the final saliency map, S.

IV. Experiment

To evaluate the performance of the proposed method, experiments were set up to investigate the efficiency of both frameworks, Hola and i-Hola, and baselines. Results were reported and discussed as follows.

A. Experimental Setup

The performance of the proposed method and baselines, including SR [2], HFT [3], SHFT [4], PQFT [5], PFDN [6], and SIG [7], is evaluated by a challenge dataset, the same as [1] used. This dataset contains seven groups with different characteristics: (i) 60 high contrast images, (ii) 100 high contrast images with complex background, (iii) 120 texture object images, (iv) 280 low contrast images, (v) 60 brightness background images, (vi) 300 luminance and shadow images, and (vii) 80 multiple salient images.

To measure our performance framework, an area under the curve (AUC), which is the most commonly used metric for saliency evaluation, is used. Its range is in between 0 to 1, such that “1” means the result of the detected saliency very similar to the ground truth, labeled by expertise, while “0” means that they are completely different.

B. Results and Discussion

Table I shows AUC performance of all test methods. It is evident that i-Hola outperformed in almost groups of dataset, especially, the overall performance of i-Hola (combined) is higher than the baseline methods, including SR [2], HFT [3], SHFT [4], PQFT [5], PFDN [6], SIG [7], and Hola (enhanced) [1] with significant improvement gain: 20.58%, 2.88%, 7.70%, 9.77%, 5.90%, 9.53%, and 0.64%, respectively.

This achievement comes from the framework improvement. The i-Hola framework (i) can provide a clearer saliency map in the stage of candidate saliency map extraction and (ii) improve decision-making in the stage of final saliency map selection. Figs. 3 and 4 visually illustrate the data flow of Hola an i-Hola frameworks, respectively. We can see that the Hola framework was failed in the final stage of saliency map selection, while i-Hola framework was successful in the final stage of saliency map selection. One reason to support this situation is applying two Hola filters in conjunction with two low-pass filters in extracting

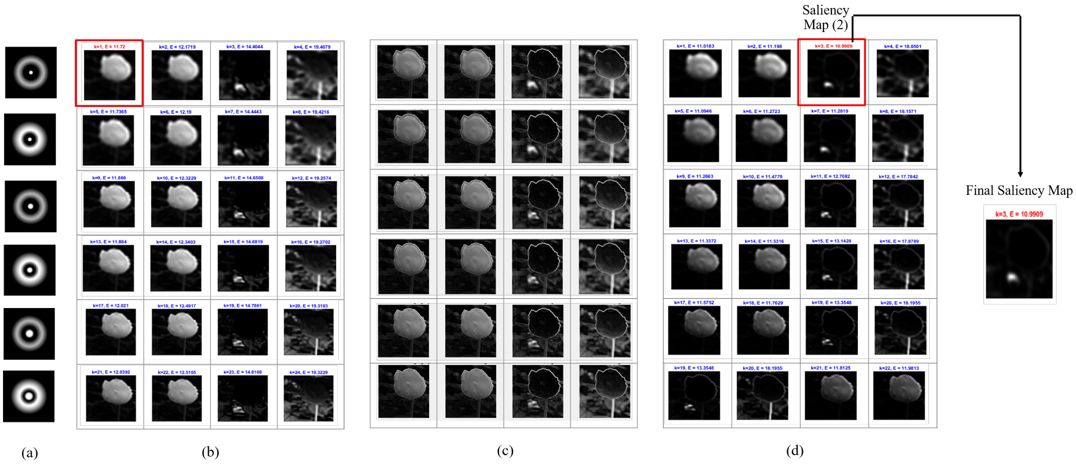

Fig. 3: An illustration of data flow of Hola framework: (a) Hola filters, (b) candidates of saliency map extraction process (1), (c) candidates of enhanced saliency maps (1), (d) candidates of saliency map extraction process (2), and final saliency map.

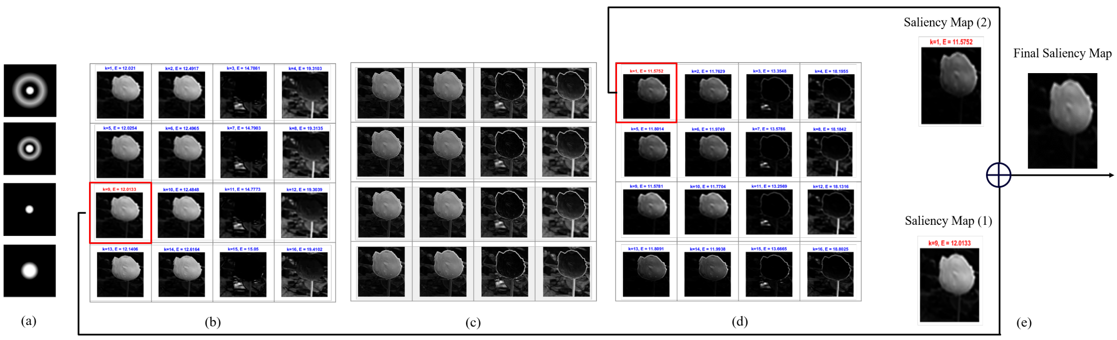

Fig. 4: An illustration of data flow of i-Hola framework: (a) i-Hola filters, (b) candidates of saliency map extraction process (1), (c) candidates of enhanced saliency map (1), (d) candidates of saliency map extraction process (2), and (e) final saliency map.

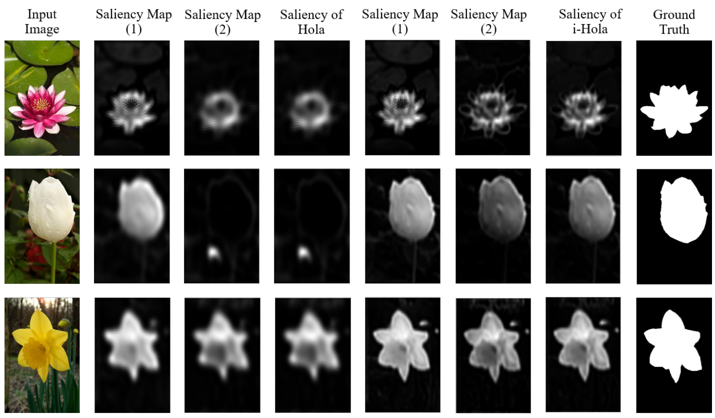

Fig. 5: A comparison of saliency maps obtained by Hola and i-Hola frameworks.

features, yielding the good quality of saliency maps. Furthermore, we modified the criterion of selecting the final saliency map by means of a pair-decision technique as defined by Eq.(1). Fig. 5 Also confirms that the i-Hola framework can provide a better result when compared with the existing Hola framework. Finally, Fig. 6 visually illustrates the significant performance improvement of i-Hola when compared with baselines.

V. Conclusion

In this paper, we have proposed an i-Hola framework for improving the quality of saliency maps to achieve (i) uniform region and (ii) sharp-edge saliency maps. Two different Hola filters and two low-pass filters were redesigned for detecting saliency in a variety of image characteristics. The implementation of this redesign filter resulted a clear saliency region. Further, to achieve sharp-edge saliency maps, the candidate saliency map enhancement was applied as a postprocessing of candidate saliency map extraction to compensate edge information as much as possible. Finally, we improved the decision-making of final saliency map selection by applying a pair-decision technique; that is, the use of the first and second rounds

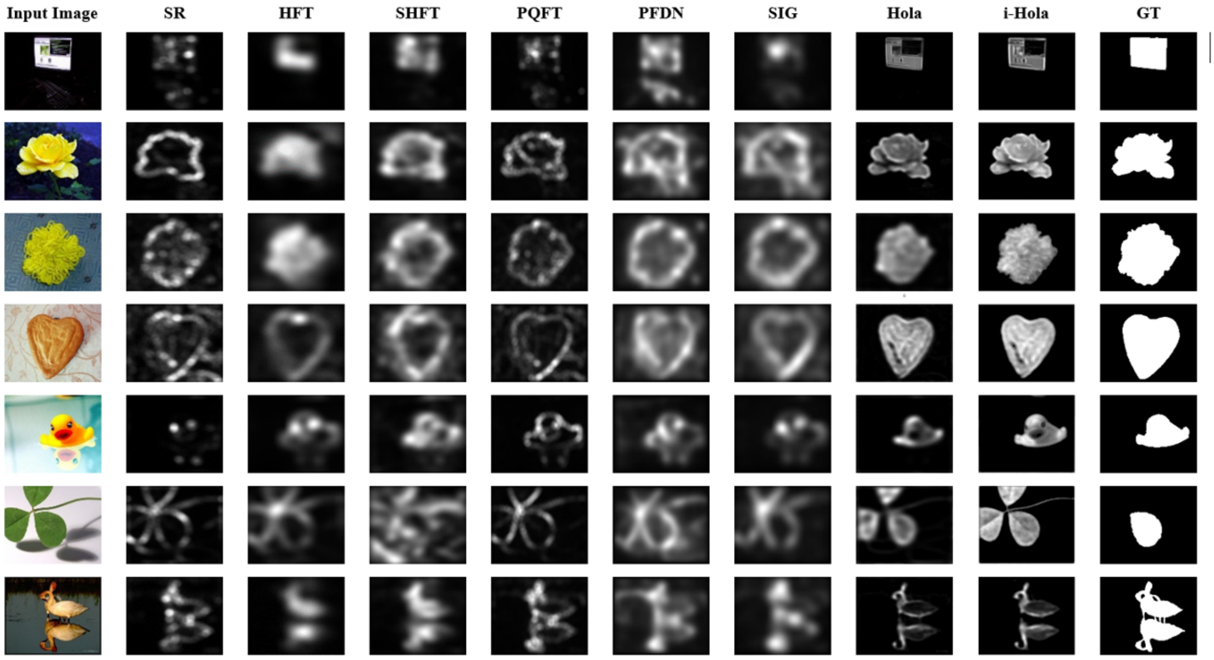

Fig. 6: Examples of saliency detection maps resulted by SR, HFT, SHFT, PQFT, PFDN, SIG, Hola, and i-Hola.

TABLE I: A comparison of AUC performance of i-Hola and baselines, including SR, HFT, SHFT, PQFT, PFDN, SIG, and Hola.

Method	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Overall
SR [2]	0.86691	0.75448	0.79184	0.76780	0.79956	0.76656	0.78350	0.77817
HFT [3]	0.93036	0.94488	0.92375	0.90384	0.87840	0.90912	0.90437	0.91204
SHFT [4]	0.92675	0.91358	0.86969	0.85698	0.81597	0.86704	0.88574	0.87123
PQFT [5]	0.90919	0.86387	0.86735	0.83815	0.87934	0.84611	0.85543	0.85481
PFDN [6]	0.92256	0.86762	0.87605	0.87646	0.91594	0.88701	0.90453	0.88610
SIG [7]	0.91058	0.84984	0.84137	0.84824	0.89841	0.85075	0.86740	0.85667
Hola [1]	**0.96400**	0.94354	**0.95600**	0.92094	0.91654	0.92878	**0.92431**	0.93235
i-Hola	0.95900	**0.95616**	0.95397	**0.93261**	**0.92594**	**0.93600**	0.92075	**0.93835**

of saliency maps to make a decision together. Based on a challenge dataset with a variety of image characteristics, the experimental results showed that our i-Hola framework outperformed all test baselines in terms of AUC performance. In future work, we look forward to extending our framework to a unified saliency segmentation.

REFERENCES

[1] D. Kakanopas and K. Worraratpanya,"An Efficient Hola Filter for Saliency Detection," 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 2020, pp.-.

[2] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, MN, pp.1-8, 2007.

[3] J. Li, M. D. Levine, X. An, X. Xu and H. He, "Visual saliency sased on scale-space analysis in the frequency domain," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.35, no.4, pp.996-1010, April. 2013.

[4] The source code is available at https://github.com/zhenglab/Saliency-Benchmark-Frequency/blob/master/Models/SHFT/

[5] C. Guo, Q. Ma and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform," in *2008 IEEE Conference Computer Vision and Pattern Recognition*, Anchorage, AK, pp.1-8, 2008.

[6] P.Bian and L. Zhang, "Visual saliency: a biologically plausible contourlet-like frequency domain approach," *Cognitive neurodynamics*, 4(3), pp.189–198, 2010.

[7] X. Hou, J. Harel and C. Koch, "Image signature: highlighting sparse salient regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.34, no.1, pp.194-201, Jan. 2012.

[8] https://mmcheng.net/msra10k/