A Universal Framework for Salient Object Detection

Jianjun Lei, *Member, IEEE*, Bingren Wang, Yuming Fang, Weisi Lin, *Fellow, IEEE*, Patrick Le Callet, Nam Ling, *Fellow, IEEE*, and Chunping Hou

***Abstract*—In this paper, we propose a novel universal framework for salient object detection, which aims to enhance the performance of any existing saliency detection method. First, rough salient regions are extracted from any existing saliency detection model with distance weighting, adaptive binarization, and morphological closing. With the superpixel segmentation, a Bayesian decision model is adopted to refine the rough saliency map to obtain a more accurate saliency map. An iterative optimization method is designed to obtain better saliency results by exploiting the characteristics of the output saliency map each time. Through the iterative optimization process, the rough saliency map is updated step by step with better and better performance until an optimal saliency map is obtained. Experimental results on the public salient object detection datasets with ground truth demonstrate the promising performance of the proposed universal framework subjectively and objectively.**

***Index Terms*—Iterative optimization, salient object detection, universal framework, visual attention.**

I. Introduction

VISUAL attention, a cognitive process to select the visually significant information from the natural scenes, is an important characteristic of the human visual system. When observing the natural image, human beings tend to focus on the salient regions rather than other irrelevant surrounding regions in the scene [1], [2]. There are two mechanisms in visual attention: bottom-up and top-down mechanisms [3]. Bottom-up attention mechanism is stimulus-driven and involuntary, while top-down attention mechanism is task-driven and voluntary. Recently, visual attention modeling has become a hot topic and has been widely used in various applications such as image retrieval [4], [5], smart video presentation [6], image compression [7], image retargeting [8], object recognition [9], sensation enhancement [10], etc. Although great progress has been made in the research area of visual attention, it still remains one of the most important and challenging issues in the fields of image analysis, pattern recognition and computer vision.

By simulating visual attention mechanism to perceive and response to visual stimuli, a saliency detection algorithm usually generates a saliency map for a given image where the value of each pixel indicates the degree this pixel stands out from the image. Early saliency detection methods mainly focus on simulating movements of human eyes to predict human fixation locations. Eye-tracking datasets are usually used to evaluate the performance of these methods. One of the classical biologically plausible models of visual attention was proposed by Itti *et al.*, which is also deemed as one of the most influential saliency detection models [11]. In [11], Itti *et al.* calculated multi-scale center-surrounding differences based on three features including intensity, color, and orientation to obtain three feature maps. These three feature maps are combined to generate the final saliency map. Similarly, Le Meur *et al*. proposed a saliency detection model based on a coherent psychovisual space from which several feature maps are combined to calculate the saliency map for images [12]. By extending the model in [11], Harel *et al.* proposed a saliency detection model based on a Markov chain on the full-connected map and a graph-based dissimilarity measure [13]. The computational model of visual attention in [14] is derived from the information maximization principle. The average transmitting information is computed by site entropy rate for saliency measure. In [15], Guo *et al.* proposed a multi-resolution spatiotemporal saliency detection model based on the principle of phase spectrum of Fourier transform (PFT). Schauerte and Stiefelhagen combined and extended several previous works on spectral saliency detection [16]. In [17], Li *et al.* proposed a saliency detection model based on statistical prior knowledge learned from millions of images.

In research of computer vision, many studies focus on detecting salient objects rather than human gaze fixations. It was suggested in [18] that early saliency has an indirect effect on attention through recognized objects, namely, human eyes are attracted by salient objects rather than a series of isolated fixation points. The salient objects generally stand out relative to their surrounding regions. The task of the salient object detection is to extract the most interesting object/objects in a scene. Accordingly, various salient object detection methods based on feature extraction from salient patches, blocks, regions or objects have been proposed. In [19], Liu *et al.* proposed a salient object detection algorithm, in which the features of multi-scale contrast, center-surround histogram, and color spatial distribution from images are extracted locally, regionally, and globally. A conditional random field (CRF) is used on these features for salient object detection in visual scenes. Jiang *et al.* utilized the

Manuscript received October 06, 2015; revised March 20, 2016 and May 30, 2016; accepted July 08, 2016. Date of publication July 18, 2016; date of current version August 12, 2016. This work was supported in part by the Natural Science Foundation of China under Grant 61271324, Grant 61520106002, Grant 91320201, Grant 61571212, and Grant 61471262. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. K. Selcuk Candan.

J. Lei, B. Wang, and C. Hou are with the School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: jjlei@tju.edu.cn; brwang@tju.edu.cn; hcp@tju.edu.cn).

Y. Fang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: fa0001ng@e.ntu.edu.sg).

W. Lin is with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: wslin@ntu.edu.sg).

P. Le Callet is with LUNAM Université de Nantes, IRCCyN UMR CNRS 6579, Polytech Nantes, Nantes Cedex 3 44306, France (e-mail: patrick.lecallet@univ-nantes.fr).

N. Ling is with the Department of Computer Engineering, Santa Clara University, Santa Clara, CA 95053 USA (e-mail: nling@scu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2016.2592325

supervised learning approach to map the regional feature vector to a saliency score and then fused the saliency scores across multiple levels to generate the saliency map [20]. In [21], Achanta *et al.* built one large-scale saliency object database and calculated the saliency map for images by DoG filters. In [22], Rahtu *et al.* applied a sliding window to measure saliency based on a rigorous statistical formulation and segmented salient object from the background through the CRF model. Gopalakrishnan *et al.* computed color saliency based on spatial distribution of color features in the image space, and orientation saliency based on the global and local behavior of the different orientations, respectively [23]. The final module in that study predicts the appropriateness of the two saliency maps and selects the one that leads to the correct identification of the salient region. Liu *et al.* proposed a nonparametric saliency detection model based on kernel density estimation (KDE) [24], [25]. The saliency map is obtained through combining saliency measures of KDE models and color likelihood measures of pixels. Taking both local and global contrast into account, Goferman *et al.* built a content-aware saliency detection model for salient region extraction [26]. Cheng *et al.* introduced a global contrast based method to produce region-based contrast maps using color statistics of the input image [27]. In [28], Gopalakrishnan *et al.* extracted salient regions through random walks on a graph modeled by color and orientation features. Both global and local saliency information are captured by random walks on a complete graph and a k-regular graph. In [29], Liu *et al.* proposed a spatiotemporal saliency model based on superpixels. The superpixel-level temporal saliency and spatial saliency are computed respectively and then derive corresponding pixel-level saliency maps. The final spatiotemporal saliency map is obtained from pixel-level temporal and spatial saliency maps through an adaptive fusion method. Perazzi *et al.* suppressed the background through uniqueness and distribution contrast of decomposed basic compactness, and then assigned saliency value to each pixel [30]. Li and Yu proposed a visual saliency model based on multiscale deep features [31]. The multiscale deep features are extracted using deep convolutional neural networks, and multiple saliency maps for different levels are aggregated to obtain the final result.

Besides the salient region extraction in the spatial domain, some studies extract salient regions in images by transform domain analysis. Hou and Zhang utilized the concept of spectral residual to build a saliency detection model [32]. They extracted the spectral residual through analyzing the log-spectrum of the image. In [33], Li *et al.* proposed to detect salient objects in images based on a scale-space analysis of the amplitude spectrum. The saliency map is generated by reconstructing the two-dimensional (2-D) signal using the original phase and the amplitude spectrum. The hypercomplex Fourier transform is used to replace the standard Fourier Transform for spectrum scale-space analysis. Fang *et al.* proposed a saliency detection model based on the human visual sensitivity and amplitude spectrum of quaternion Foutier transform (QFT) [34]. The saliency value is calculated by differences between the QFT amplitude spectrum of patches and the visual impacts for these differences determined by the human visual sensitivity. Imamoglu *et al.* detected salient object utilizing low-level features obtained from the wavelet transform domain. Various feature maps are generated by IWT in various scales and used to form the final saliency map [35]. In [36], Fang *et al.* calculated saliency in the compressed domain. The features of color, intensity and texture are extracted from DCT coefficients in the JPEG bit-stream. The saliency of each DCT block is measured through Hausdorff distance calculation and feature map fusion.

As introduced above, there is a number of saliency detection models proposed for various multimedia applications. Different from existing saliency detection studies which aim to design a new method for saliency detection, we propose a universal framework to improve the performance of any existing saliency detection model based on the rough saliency map from the saliency detection model. Given any existing saliency detection model, we first extract the rough saliency map from an input image. By using the rough saliency map as the prior knowledge, a Bayesian decision model is built to compute a more accurate saliency map for the image. An iterative optimization process is designed to calculate the better saliency map for salient object detection at each time. Through the iterative optimization process, the rough saliency map is updated step by step with better performance until an optimal saliency map is obtained. Experimental results demonstrate that the proposed universal framework can obtain much performance improvement over the existing saliency detection models.

The rest of the paper is organized as follows. Section II introduces the proposed method in detail. Experimental results are presented in Section III. Finally, Section IV gives the conclusion.

II. Proposed Method

The block diagram of the proposed universal framework is presented in Fig. 1. First, with the operations of distance weighting, adaptive binarization, and morphological closing, the rough salient regions are extracted in the saliency map from any saliency detection method. Then the rough salient region and the superpixels are utilized to generate a more accurate saliency map based on Bayesian decision theory. Finally, an iterative optimization process is designed to calculate the optimal saliency map for the input image.

A. Rough Salient Region Extraction

In this section, we introduce the process to extract the rough salient region, which is the basic process of our approach. The saliency map calculated by any existing method is utilized as a prior saliency map to extract the rough salient region.

In the proposed universal framework, the more accurate the rough salient region is, the better it is for the saliency computation under the Bayesian decision model. The rough salient region with high accuracy is highly desirable. Here, we use distance weighting, adaptive binarization, and morphological closing operation to extract the rough salient region based on the prior saliency map of an input image.

It is well known that human beings tend to pay more attention to the center of the scene than the surrounding regions when viewing visual scenes on a display [37]. Based on this prior

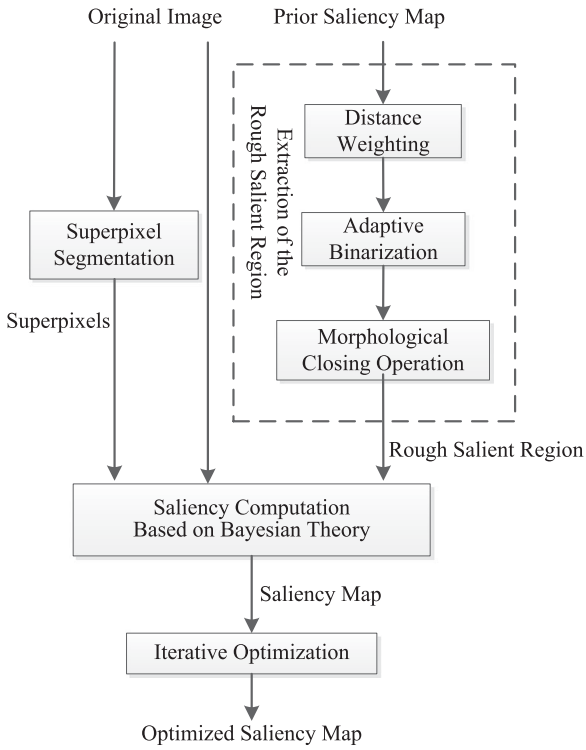


Fig. 1. Block diagram of the proposed universal framework.

knowledge, we first conduct distance weighting on the prior saliency maps according to the distance from the image center. The distance weighting is calculated as

$$I(x,y) = I_0(x,y) \cdot \exp\left(\frac{(\alpha - r)^2}{\alpha^2}\right) \tag{1}$$

where $I_0(x,y)$ indicates the pixel values of the prior saliency maps; $I(x,y)$ indicates the pixel values of the saliency maps after distance weighting; α denotes the distance between the image center and the farthest pixels, expressed as

$$\alpha = \sqrt{\left(\frac{w}{2}\right)^2 + \left(\frac{h}{2}\right)^2} \tag{2}$$

where w and h are the width and height of the image, respectively. r represents the Euclidean distance between the point and the image center, expressed as

$$r = \sqrt{\left(x - \frac{w}{2}\right)^2 + \left(y - \frac{h}{2}\right)^2}. \tag{3}$$

After the distance weighting operation, the normalization operation is conducted on the saliency maps.

Otsu method is a simple and efficient algorithm to convert gray image into binary image with an adaptive threshold [38]. In this study, the adaptive binarization is implemented using Otsu method to select the threshold and segment the saliency maps into binary maps. A binary map is obtained in which the regions with value 1 are salient while the regions with value 0 represent the background.

As the binary maps may include incoherent regions with small sizes, the morphological closing operation is implemented on the binary maps. In our experiment, the radius of the structuring element in the morphological closing operation is empirically set as 10 pixels. Accordingly, a complete rough salient region is extracted. An example of the rough salient region extraction is shown in Fig. 2. The prior saliency map in Fig. 2(b) is obtained from the salient object detection method of RC [27]. The extracted rough salient region is shown in Fig. 2(c).

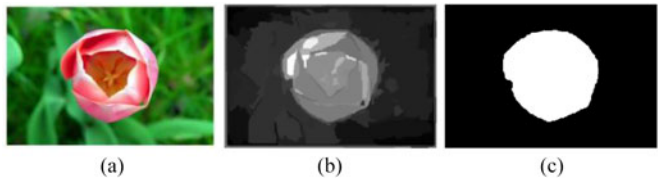


Fig. 2. Example of the rough salient region extraction. (a) Original image. (b) Prior saliency map. (c) Binary map with the rough salient region.

B. Saliency Computation Based on Bayesian Decision Theory

Based on the Bayesian decision theory, the saliency computation can be formulated as estimating the posterior probability at each pixel x in the image [39]. We can compute the prior probability and the observed conditional probability at each pixel x of the image to generate the posterior probability distribution, namely, the final saliency map.

To formalize the saliency computation of the proposed universal framework based on the Bayesian decision theory, let x denote a pixel in the image, and the binary variable C denote whether or not a pixel belongs to the salient object. The formulation of saliency computation for the universal framework is expressed as follows:

$$\begin{aligned} S(x) &= p(C{=}1|x) \\ &= \frac{p(C=1)\,p(x|C=1)}{p(C=1)\,p(x|C=1) + p(C=0)\,p(x|C=0)} \end{aligned} \tag{4}$$

where $C{=}1$ denotes the pixel belonging to the salient object, and $C{=}0$ denotes the pixel belonging to the background; $p(C{=}1|x)$ indicates the posterior probability at each pixel x in the image; $p(C=1)$ and $p(C=0)$ represent the prior probability of being salient or belonging to the background at each pixel x, respectively. The relationship of these two components is as follows: $p(C=0) = 1 - p(C=1)$. $p(x|C=1)$ and $p(x|C=0)$ indicate observed conditional probability distribution obtained from the rough salient region and the background region, respectively.

Superpixel algorithms have been widely used to segment texture image into several regions. Among the superpixel algorithms, SLIC [40] is an algorithm that can cluster pixels in the combined 5-D color and image plane space to efficiently generate compact and nearly uniform superpixels with low computational complexity. Moreover, it can obtain excellent boundary adherence. SLICO, the zero parameter version of the SLIC algorithm, can adaptively choose the compactness parameter for each superpixel differently [40], [41]. In this study, we use SLICO algorithm to segment the original image into several

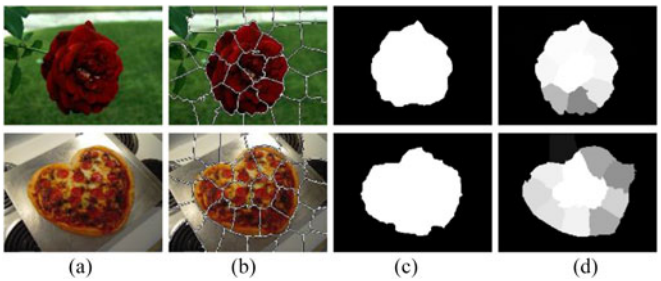


Fig. 3. Examples of prior probability distribution generation. (a) Original image. (b) Superpixel map after SLICO segmentation. (c) Binary map with the rough salient region. (d) Distribution map of prior probability.

regions. Examples for superpixels images after SLICO segmentation are shown in Fig. 3(b).

With the superpixels and the extracted rough salient region, we can compute the prior probability distribution. The rough salient region is considered as the approximate position of the salient object, and then the prior probability of each superpixel can be measured by the number of pixels in the superpixel belonging to the rough salient region. The prior probability of saliency for pixels of each region is defined as

$$p\left(C=1\right)=\frac{N\left(SP\cap RSR\right)}{N\left(SP\right)} \tag{5}$$

where SP represents a superpixel after superpixel segmentation, and RSR denotes the rough salient region. Then $N(SP\cap RSR)$ indicates the number of pixels in a superpixel belonging to the rough salient region, and $N(SP)$ indicates the whole number of pixels in the superpixel.

Through the above calculation, the prior probability distribution of saliency for pixels of each superpixel can be obtained according to the degree with which the region overlapped with the rough salient region. Examples of prior probability distribution generation are shown in Fig. 3. Fig. 3(c) shows binary maps with the rough salient region extracted from the prior saliency maps of RC. The prior probability distribution maps are shown in Fig. 3(d).

We then compute the observed conditional probability utilizing the rough salient region. The saliency map can be divided into two components: the rough salient region RSR and the background region BKG. Intuitively, the pixels in RSR incline to being salient while those in BKG are more likely to belong to the background.

As CIELab color space is consistent with human visual perception system, the input original color image in RGB color space is transformed into the CIELab color space, in which the luminance channel and the two chrominance channels are well decorrelated [42]. The color features of the regions RSR and BKG are calculated as three components L*, a* and b* for each pixel x. The observed conditional probability for the region RSR is computed as

$$p\left(x|C=1\right)=\prod_{I\in\{L^*,a^*,b^*\}}\frac{C_{RSR}\left(I\left(x\right)\right)}{C_{RSR}} \tag{6}$$

and the observed conditional probability for the region BKG is computed as

$$p\left(x|C=0\right)=\prod_{I\in\{L^*,a^*,b^*\}}\frac{C_{BKG}\left(I\left(x\right)\right)}{C_{BKG}} \tag{7}$$

where C_{RSR} represents the count of pixels in region RSR, and C_{BKG} represents the count of pixels in region BKG. $C_{RSR}(I(x))$ indicates the region RSR containing the number of pixels with the same color value of pixel x in I channel, while $C_{BKG}(I(x))$ indicates the region BKG containing the number of pixels with the same color value of pixel x in I channel.

In brief, we first compute the prior probability of a superpixel using (5). Then, the observed conditional probability is computed according to (6) and (7). At last, the output saliency map is generated based on the posterior probability computation using the Bayesian decision theory [see (4)].

C. Iterative Optimization Process

To obtain better saliency results, we use the obtained saliency map from (4) as the prior saliency map to generate the new saliency map. An iterative optimization process is designed to obtain more satisfying saliency maps.

For simplicity, the prior saliency map calculated from a saliency detection method is expressed as I_0, and the output saliency map from (4) is expressed as UM_0. The output saliency map from the first iterative optimization process is expressed as UM_1, the output saliency map of the second iterative optimization process is expressed as UM_2, and so on.

In the first iterative optimization process, the saliency maps I_0 and UM_0 are utilized to extract the new rough salient region. As described in Section II-A, we successively conduct the steps of distance weighting and adaptive binarization on saliency maps I_0 and UM_0, respectively. Thus, two binary maps are calculated based on I_0 and UM_0 respectively. These two binary maps are multiplied to extract their intersection, which can largely eliminate the regions that do not belong to the salient object. Accordingly, a more accurate rough salient region can be obtained. Similarly, the morphological closing operation is implemented on the above intersection to exclude the incoherent regions with small sizes. The output saliency map of the first iterative optimization process UM_1 is then generated through saliency computation based on the Bayesian decision theory. The overview of the first iterative optimization process is expressed as Fig. 4.

In the second iterative optimization process, the saliency maps I_0, UM_0 and UM_1 are utilized to extract the rough salient region. Similar as the first iterative optimization process, three binary maps are computed based on I_0, UM_0 and UM_1 respectively through the steps of distance weighting and adaptive binarization. These three binary maps are multiplied to extract their intersection. The morphological closing operation is conducted to the intersection region to obtain the new rough salient region. Based on the updated rough salient region, the output saliency map of the second iterative optimization process UM_2 is then generated through computing saliency based on the Bayesian decision theory.

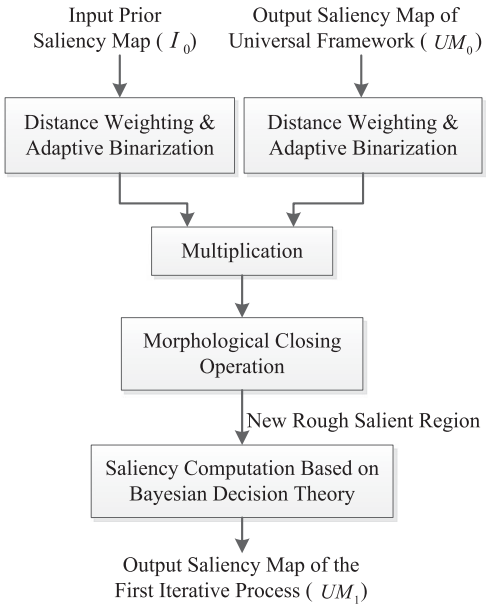


Fig. 4. Overview of the first iterative process.

Similarly, the next optimization iterations are continually conducted with the previous iterative optimization results. The performance of the output saliency map can be improved as the iterative optimization process is carried on. Iterative optimization processes will be continued until the saliency results are stable.

A termination strategy is designed for the iterative optimization process to obtain the final saliency map. The judgment criterion of the strategy takes the average difference between the pixel values of the output saliency maps. When the difference between the output saliency map of the current iterative process and that of the previous iteration reaches to a certain degree, the iterative optimization process will stop. Then the output saliency map of the last iteration is considered as the final saliency map, which is expressed as UM_F.

The judgment criterion of the termination strategy for the iterative optimization process can be formulated as follows:

$$TS = \frac{1}{w \cdot h}\left(\sum_{x=1}^{w}\sum_{y=1}^{h}\left|UM_m\,(x,y) - UM_{m-1}\,(x,y)\right|\right) \quad (8)$$

where TS is the average difference of pixel values; $UM_m(x,y)$ and $UM_{m-1}(x,y)$ denote the values of the saliency maps from the *m*th and (*m*-1)th iterative optimization processes, respectively; w and h represent the width and height of the saliency map, respectively.

The average difference of pixel values, TS, is computed after the output saliency maps are obtained for each iterative optimization process. If the average difference of pixel values is less than a given threshold T, the iterative optimization process will stop. Furthermore, for the case of non-convergence, the iterative process will stop when the maximum number of iteration *M* is reached, even though the average difference does not reach the given threshold. Visual samples of our universal framework are shown in Fig. 5. As shown in Fig. 5(d), the rough salient region of RC is extracted based on the prior saliency map using some operations, including distance weighting, adaptive binarization, and morphological closing operation. The performance is significantly improved after saliency computation based on Bayesian decision theory, as shown in Fig. 5(e). Through the iterative optimization process, more accurate and intact salient regions are obtained, as shown in Fig. 5(f).

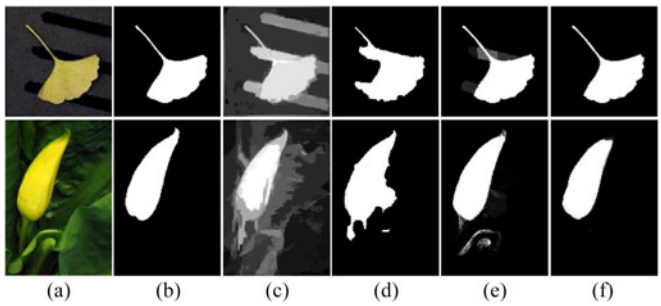


Fig. 5. Visual samples from the proposed universal framework. (a) Original image. (b) Ground truth. (c) Original input saliency map of RC. (d) Binary map with the rough salient region. (e) The output saliency map of the proposed universal framework. (f) Saliency map after the iterative optimization process.

III. Experimental Results

We have evaluated the results of the proposed method on the publicly available dataset MSRA-1000 [21] including 1000 natural images with their corresponding ground truth masks produced by human-marked labeling for salient objects, and a more complex dataset, ECSSD [43], with 1000 images and the corresponding ground truth masks.

A. Performance Evaluation of the Proposed Universal Framework

We first test the proposed universal framework on the salient object dataset MSRA-1000. Several state-of-the-art approaches, including SR [32], FT [21], LC [44], HC [27], RC [27], SF [30], and PCA [45] are used to conduct the comparison experiment in this study.

In our experiment, the image is segmented into 30 regions in advance by SLICO algorithm for our universal framework. For several state-of-the-art methods, including SR, FT, LC, HC, RC, and SF, we compare the saliency maps with and without the implementation of the proposed universal framework.

We provide some visual samples in Fig. 6, in which both saliency maps I_0 and UM_0 from these state-of-the-art methods and the proposed universal framework are shown for visual comparison. From the visual comparison, we can see that the performance of the saliency maps obtained from all these six state-of-the-art methods has been improved after the implementation of the proposed universal framework.

In order to objectively evaluate the performance of the proposed universal framework, the measures of precision and recall are used for comparison, similar as previous works. The precision-recall curves are also provided, in which higher precision values and recall values indicate better performance with fixed threshold value to the saliency map. The precision and

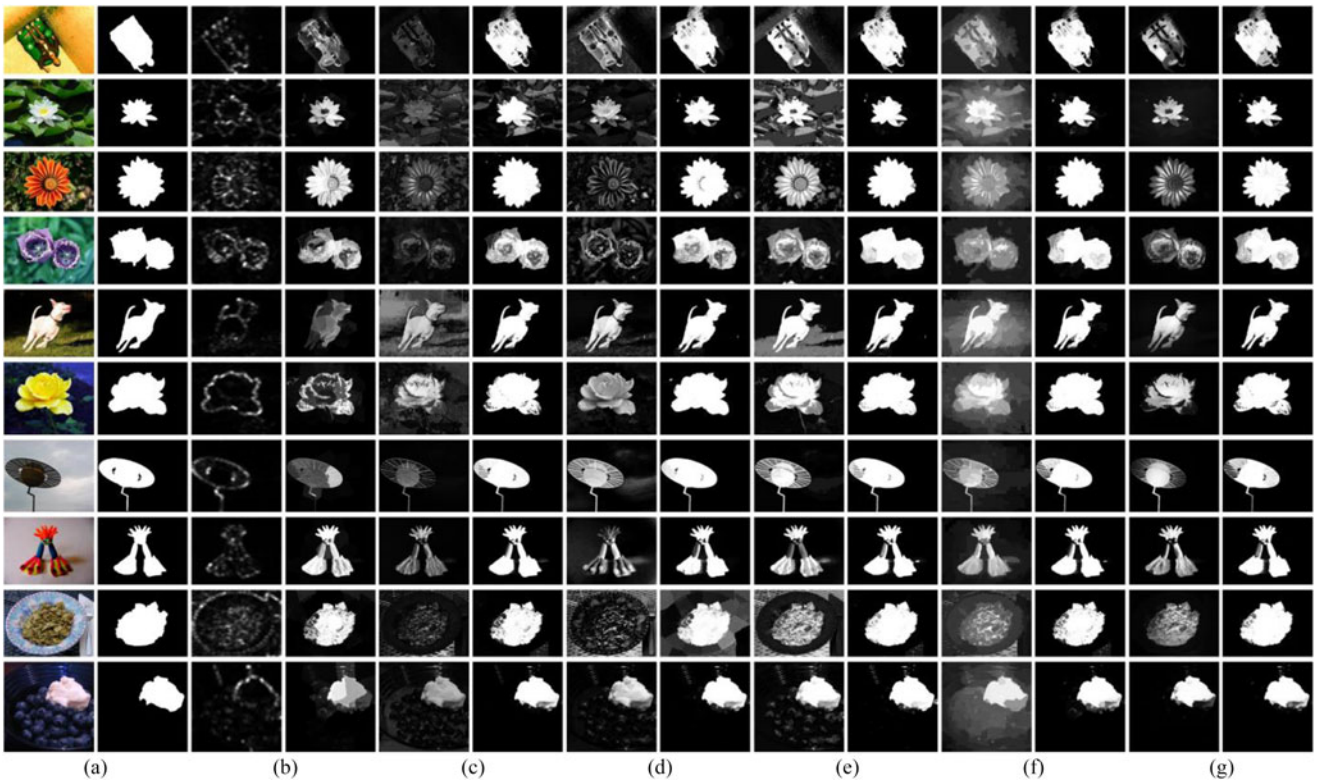


Fig. 6. Visual samples of salient object detection results. (a) Input original images and corresponding ground truths. (b)–(g) The saliency maps from different methods and their corresponding saliency maps from the proposed universal framework for (b) SR, (c) FT, (d) LC, (e) HC, (f) RC, and (g) SF.

recall values are defined as

$$precision = \frac{\sum S \cap G}{\sum S} \tag{9}$$

$$recall = \frac{\sum S \cap G}{\sum G} \tag{10}$$

where S denotes the binary saliency pixels, G denotes the ground truth pixels and Σ refers to the sum of all pixels. We use a series of fixed integer thresholds from 0 to 255 to obtain 256 binary salient object masks for each saliency map. With each threshold, the precision and recall values are averaged, and the precision-recall curve plots the average precision values against the average recall values as shown in Fig. 7.

It can be seen from Fig. 7 that the precision-recall curves of the results after the implementation of our universal framework is much higher than those from the original saliency maps of the six state-of-the-art methods. Accordingly, we can conclude that the quality of the saliency maps after processing of the proposed universal framework is generally quite better than that of the original saliency maps obtained from the six state-of-the-art methods.

In addition, we further utilize the F-measure to objectively evaluate the quality of the proposed universal framework, which is defined as

$$F_\beta = \frac{\left(1+\beta^2\right) \cdot precision \cdot recall}{\beta^2 \cdot precision + recall} \tag{11}$$

where β^2 is set to 0.3 in our experiments to weight precision more than recall as suggested in [21], and we use the image saliency dependent adaptive threshold proposed by [21] to generate binary maps of salient object from saliency maps.

Fig. 8 shows the precision, recall, and F-measure of the results of the seven state-of-the-art methods and corresponding results after processing of our universal framework on MSRA-1000 dataset. It is obvious that the F-measure and recall values increase after the implementation of the proposed universal framework for all seven existing methods. High recall indicates that it benefits for salient object segmentation after implementation of the universal framework. Please note that the precision does not increase after implementation of the universal framework for SF, but the recall and F-measure values are both much higher than those from the original methods.

In addition, we use a more complex dataset ECSSD with 1000 images and the corresponding ground truth masks to further evaluate the performance of the proposed universal framework. Fig. 9 shows the precision, recall, and F-measure of several existing saliency detection methods and corresponding results with the proposed universal framework on ECSSD dataset. It can be observed form the figure that the precision, recall, and F-measure values are all improved after the implementation of the proposed universal framework.

Considering the flaws of the above evaluation measures, we further use the weighted F_β^w-measure [46] to evaluate the proposed universal framework. The weighted F_β^w-measure is

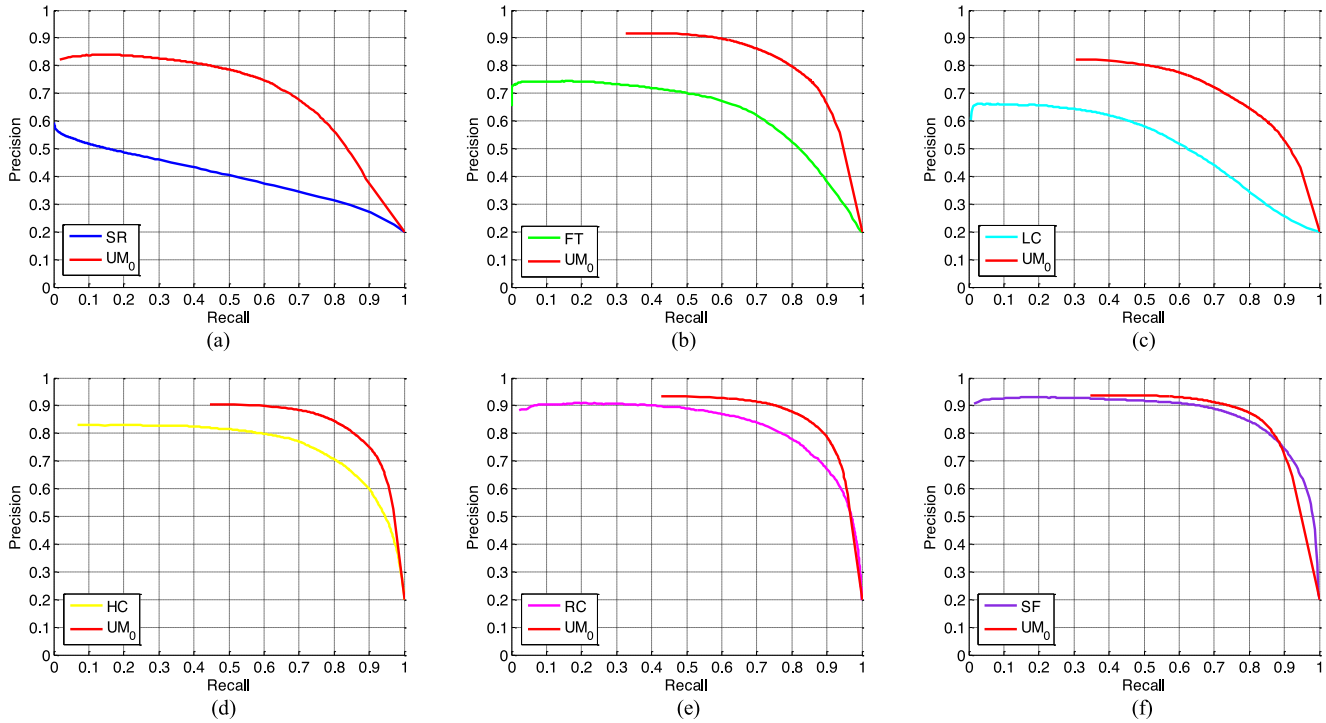


Fig. 7. Precision-recall curves generated using original saliency maps of six state-of-the-art methods and their corresponding saliency maps with the implementation of the universal framework. Results for (a) SR, (b) FT, (c) LC, (d) HC, (e) RC, and (f) SF.

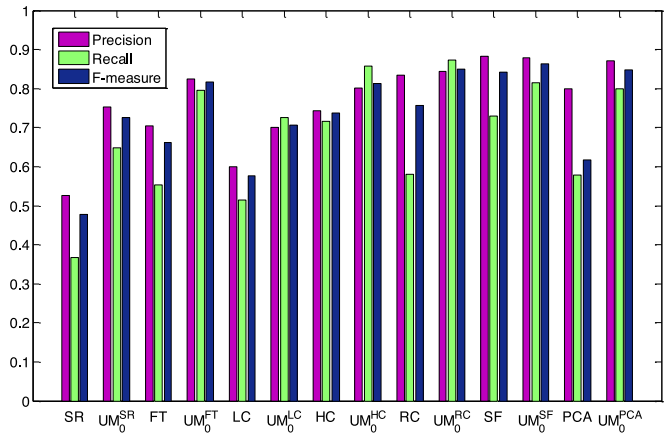


Fig. 8. Precision, recall, and F-measure results achieved by using original saliency maps of several existing saliency detection methods and those corresponding saliency maps after the implementation of the universal framework on MSRA-1000 dataset.

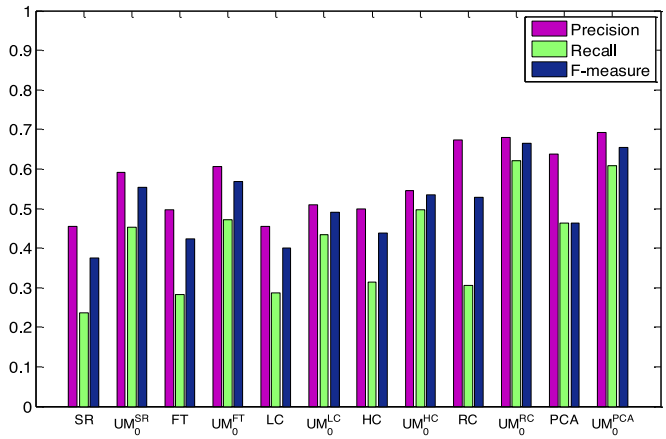


Fig. 9. Precision, recall, and F-measure results achieved by using original saliency maps of several existing saliency detection methods and those corresponding saliency maps after the implementation of the universal framework on ECSSD dataset.

defined as

$$F_\beta^w = \frac{\left(1+\beta^2\right)\cdot precision^w \cdot recall^w}{\beta^2 \cdot precision^w + recall^w}. \tag{12}$$

The results are shown in Figs. 10 and 11 for MSRA-1000 dataset and ECSSD dataset, respectively. It can be observed from the figures that the F_β^w-measure values increase with the proposed universal framework.

Overall, the proposed universal framework is able to improve the performance of existing salient detection methods.

B. Performance Evaluation of the Proposed Iterative Optimization Process

In order to evaluate the performance of the proposed iterative optimization process, the iterative optimizations with different times are implemented for the seven state-of-the-art methods under the termination strategy. The experiments are conducted on the basis of the results from our model, UM_0, and terminated after obtaining the final saliency maps, UM_F. The threshold of the termination strategy T is set to 1 by considering both the computational complexity and the performance of saliency

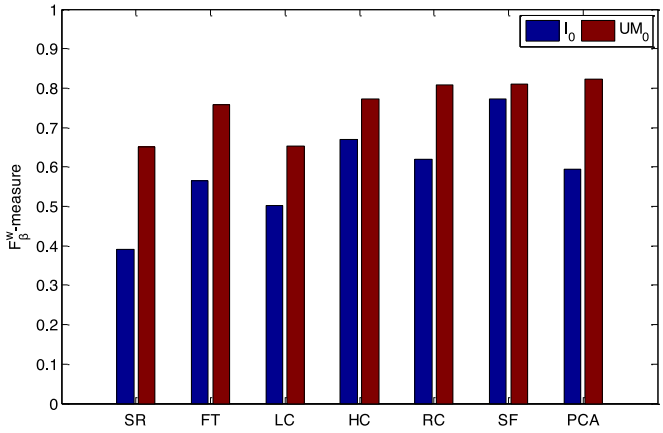


Fig. 10. Weighted F_β^w-measure results achieved by using original saliency maps of several existing saliency detection methods and those corresponding saliency maps after the implementation of the universal framework on MSRA-1000 dataset.

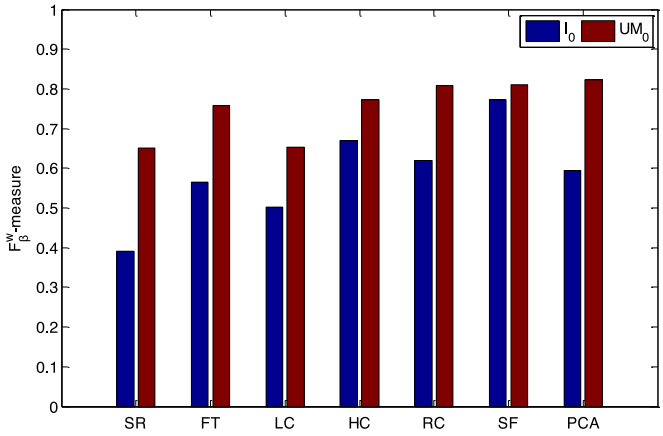


Fig. 11. Weighted F_β^w-measure results achieved by using original saliency maps of several existing saliency detection methods and those corresponding saliency maps after the implementation of the universal framework on ECSSD dataset.

maps. The maximum number of iteration *M* is set to 6. In order to ensure the overall performance, the first three iterations are conducted necessarily. In the experiments, most saliency detection methods can obtain the final saliency maps with three times of iterative optimization process.

We provide some visual samples of the results from the seven state-of-the-art methods in Figs. 12 and 13. It can be seen from Fig. 12 that the performance of the saliency maps is better and better after each iterative optimization process, which demonstrates that the iterative optimization process of our universal framework is able to further improve the performance of saliency maps obtained from other saliency detection methods.

In addition, we further utilize the F-measure to objectively evaluate the performance of the proposed iterative optimization process. As shown in Fig. 14, the F-measure values of each iterative optimization process for the seven state-of-the-art saliency detection methods are plotted in the light of corresponding saliency maps. We can see from the figure that the F-measure values are much higher at the initial stage of the

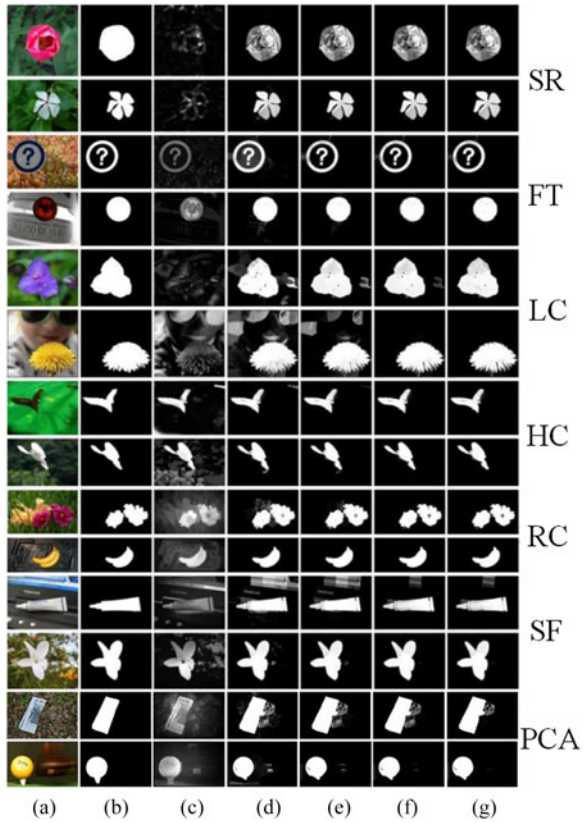


Fig. 12. Visual comparison of some results from the proposed iterative optimization process for seven state-of-the-art methods. (a) Input original images. (b) Ground truths. (c) Input original saliency maps of the seven state-of-the-art methods. (d) Saliency maps UM_0 after implementation of our universal framework. (e) Saliency maps UM_1 after the first iterative optimization process. (f) Saliency maps UM_2 after the second iterative optimization process. (g) Saliency maps UM_F after the last iterative optimization process.

iterative optimization process and incline to be slightly improved or stable as the iterative optimization process continues.

For input images whose initial saliency map UM_0 is satisfactory enough, there is no much performance improvement from the next iterative optimization processes of the proposed universal framework, as shown in Fig. 13. It can be seen from the Fig. 13(b)–(d) that the saliency map from universal framework is much better than the original saliency map. However, there is no much improvement for the saliency results with the iterative optimization operation compared with those without iterative optimization operation. From all the saliency results in Fig. 13, we can see that the performance of the output saliency maps tends to be stable as the iterative optimization process continues.

C. Detailed Analysis

In this section, we first give a detailed analysis to show how the proposed framework improves a bad result from the original saliency map. The detailed results from the proposed universal framework for SR are shown in Fig. 15. As shown in Fig. 15(c), there are blurred boundaries and incomplete details in the original saliency map of SR. However, the result is significantly improved with the proposed universal framework, as

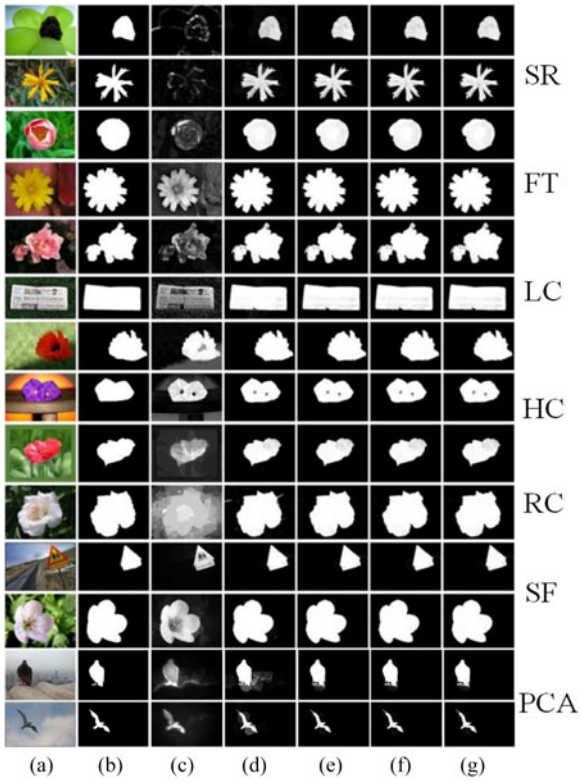


Fig. 13. Visual samples of saliency results from the proposed iterative optimization process for seven state-of-the-art methods. (a) Input original images. (b) Ground truths. (c) Input original saliency maps of the seven state-of-the-art methods. (d) Saliency maps UM_0 after implementation of our universal framework. (e) Saliency maps UM_1 after the first iterative optimization process. (f) Saliency maps UM_2 after the second iterative optimization process. (g) Saliency maps UM_F after the last iterative optimization process.

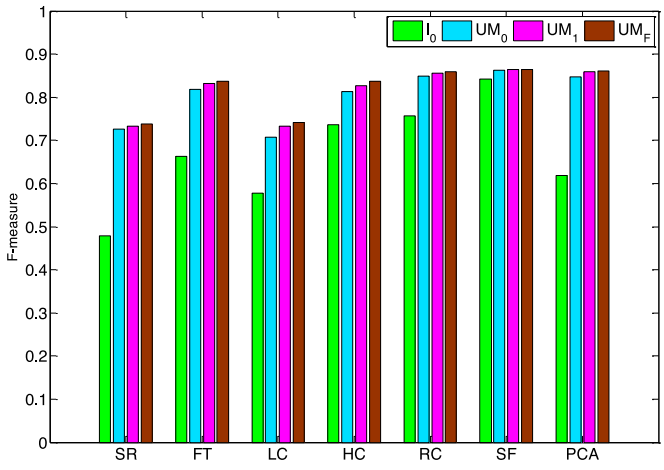


Fig. 14. F-measure values achieved by using original saliency maps of seven state-of-the-art methods and those corresponding saliency maps after the implementation of iterative optimization process of the universal framework.

shown in Fig. 15(i). Through the distance weighting, adaptive binarization, and morphological closing operations, the rough salient regions become more accurate and intact, as shown in Fig. 15(e). The prior probability is computed based on the rough saliency map [see Fig. 15(e)] and the segmented superpixels

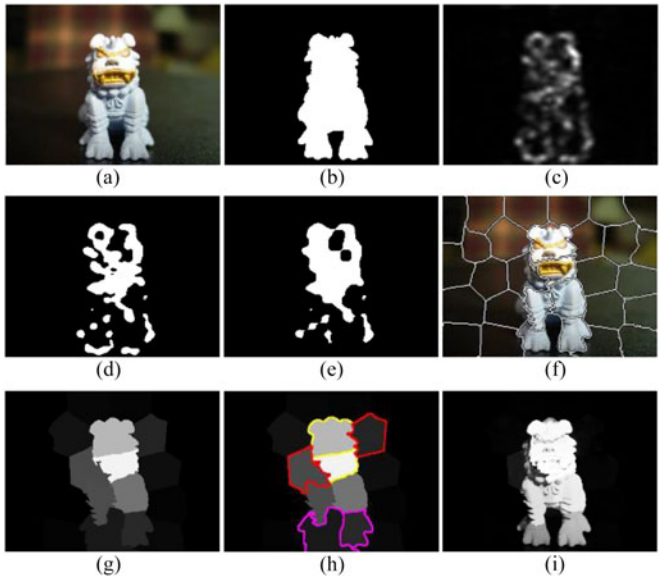


Fig. 15. Details of the proposed universal framework for the initial saliency map of SR. (a) Original image. (b) Ground truth. (c) Original input saliency map of SR. (d) Binary map after distance weighting and adaptive binarization. (e) Binary map after morphological closing operation. (f) Superpixel map. (g) Prior probability map. (h) Prior probability map with labels. (i) The output saliency map of the proposed universal framework.

[see Fig. 15(f)]. If a superpixel contains more salient pixels, the superpixel will have a higher prior probability, as shown in Fig. 15(g) and (h). For example, the prior probabilities of the two superpixels marked with yellow lines in Fig. 15(h) are 0.7531 and 1, respectively. Similarly, the salient regions near the image boundary achieve a non-zero prior probability value by SLICO method and the rough salient regions. For example, the prior probabilities of the two superpixels marked with magenta lines in Fig. 15(h) are 0.1687, and 0.0576, respectively. However, the superpixel segmentation results from SLICO method may not always be satisfactory. In some cases, the extraction of the salient object contour is not accurate, and some background regions are incorrectly included in salient regions. For example, the two superpixels marked with red lines in Fig. 15(h) do not have good superpixel contour, resulting in false prior probabilities (with values of 0.3045 and 0.1481, respectively).

In addition, the observed conditional probability by using the rough salient region and the number of pixels in color space also plays an important role in the proposed framework. Since the observed conditional probabilities are with the ability to distinguish the color difference between different regions, it can eliminate the errors in prior probability map. After the computation based on the Bayesian decision framework, the right superpixel marked with the red line is almost completely suppressed, and the left superpixel marked with the red line is also largely suppressed. Most of those regions should be distinguished from the salient object. Furthermore, the output saliency map is improved through the iterative optimization process. All of these operations guarantee that our framework can improve the performance of the existing saliency detection method.

In a word, for the bad results from the original input saliency maps, several components of the proposed framework play

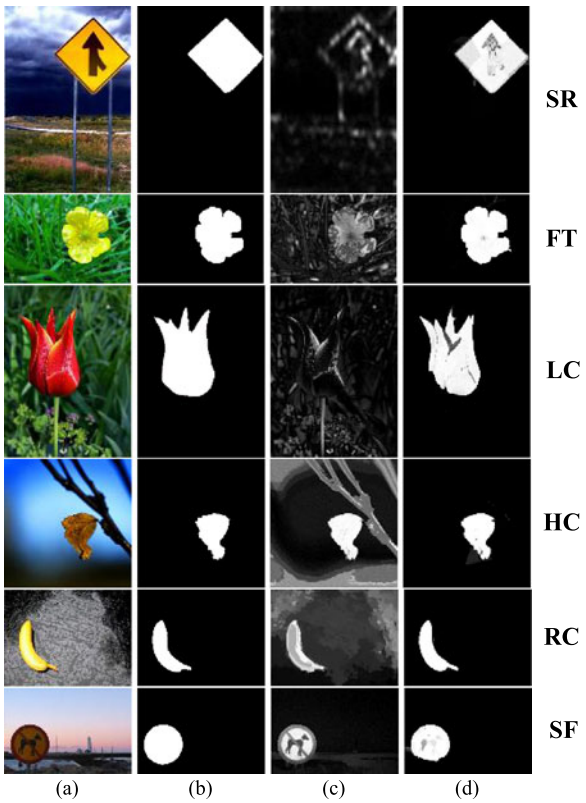


Fig. 16. Visual samples where the salient object is away from the center of the image. (a) Original images. (b) Ground truths. (c) Original input saliency maps. (d) Saliency maps of our method.

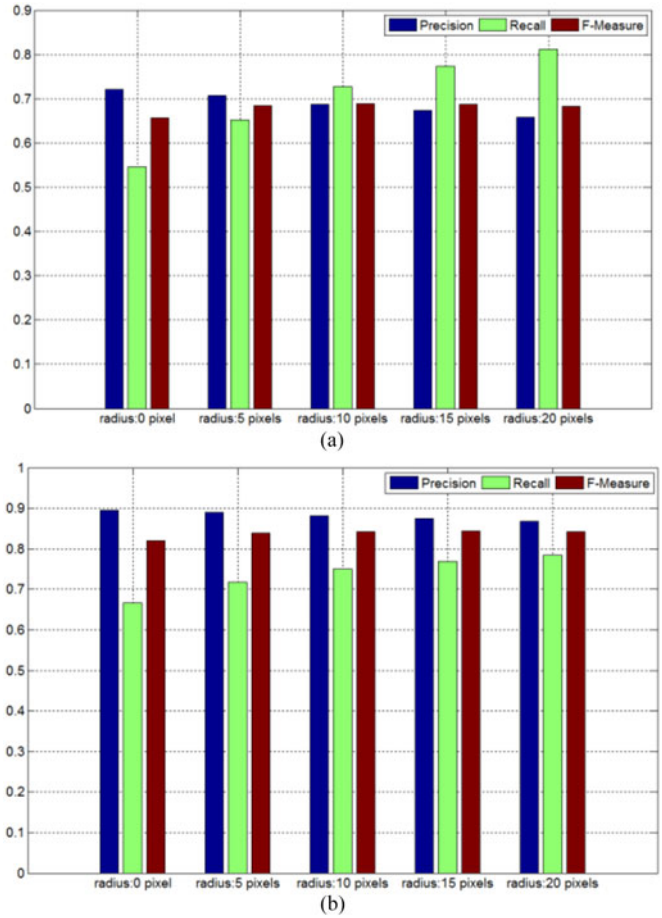


Fig. 17. Influence of morphological closing radius for UM0 using different prior saliency maps. (a) LC. (b) SF.

important roles to obtain the good saliency results, including the rough salient region extraction, SLICO segmentation, and the observed conditional probability.

We also conduct an exhaustive experiment with the salient regions away from the image center, as shown in Fig. 16. From the figure, it can be observed that the proposed universal framework can achieve a more accurate result no matter whether the salient object is located near the center of the image or not. It indicates that our universal framework is robust to the spatial position of salient objects.

Furthermore, we analyze the influence of morphological closing, and the experimental results are shown in Fig. 17. From these results, it can be observed that the performance with morphological closing is improved obviously. It is reasonable to introduce morphological closing in the proposed framework. Meanwhile, the effects with different radiuses (5, 10, 15, and 20) are tested in the comparative experiments. Considering that the performance is not sensitive to the radius, we empirically set the value of radius to be 10.

D. Computational Complexity and Discussions

To demonstrate the efficiency of the proposed method, we conduct the experiments of computational complexity by using the existing saliency detection methods and the proposed universal framework as well as the iterative optimization process. The experiments are performed on a Quad Core 3.7 GHz machine with 16 GB RAM. SR [32], FT [21], LC [44], HC [27], and RC [27], are implemented in C++. For SF [30], we do not provide its computational complexity since we could not get the source code of this model. However, the authors of SF indicate that the computational complexity of SF is similar to that of RC in [30]. PCA [45], the proposed universal framework and the iterative optimization process are implemented in Matlab. We provide the computational complexity of different models in Table I. It can be

TABLE I
COMPUTATIONAL COMPLEXITY ANALYSIS BASED ON THE AVERAGE RUNNING TIME

Time(s)	SR	FT	LC	HC	RC	SF	PCA
I_0	0.003	0.025	0.004	0.116	0.971	N/A	1.885
UM_0	0.318	0.329	0.338	0.340	0.334	0.328	0.324
UM_1	0.342	0.371	0.365	0.359	0.374	0.353	0.364
UM_2	0.378	0.409	0.405	0.411	0.403	0.412	0.405

The row of I_0 indicates the average time taken by each existing saliency detection method. The row of UM_0 indicates the average time taken by proposed universal framework. The row of UM_1 indicates the average time taken by the first iterative optimization process. The row of UM_2 indicates the average time taken by the second iterative optimization process.

observed from the table that the computational complexity of the proposed universal framework is low. The time cost of the proposed framework is less than 0.5 second to process one image at each iteration. In particular, under the same platform and computational environment, the total time cost with three times of iteration by the proposed framework is even less than that of PCA. In addition, the running time of the proposed framework through three iterations is almost the same with RC, even though the proposed framework is implemented in Matlab and RC is implemented in C++. However, it is worthwhile to mention that, the proposed framework may lead to more than three iterations to converge.

Different from existing saliency detection studies which aim to design a new method for saliency detection, we propose a universal framework to improve the performance of any existing saliency detection model. The rough saliency map obtained from existing saliency detection model is regarded as the prior knowledge, and a Bayesian decision model is built to compute a more accurate saliency map for the image. Furthermore, an iterative optimization process is designed to calculate a better saliency map at each time. It is worthwhile to mention that, our work is a primary attempt to improve the existing saliency map, and the proposed framework is an open system where other existing methods can be included easily. In the future, we would like to optimize the proposed framework by updating the operations including segmentation, rough salient region extraction, and iterative optimization mechanism.

IV. Conclusion

In this paper, we have presented a universal framework for salient object detection, which can be used to improve the performance of existing saliency detection methods. The proposed universal framework uses the saliency map of any saliency detection method to extract a rough salient region, and compute a more accurate saliency map based on this rough salient region under Bayesian decision theory. An iterative optimization process is designed to exploit the output saliency map with more accurate rough salient region. Extensive experimental results on the public salient object detection datasets demonstrate the promising performance of the proposed universal framework subjectively and objectively.

References

[1] J. Li, D. Xu, and W. Gao, "Removing label ambiguity in learning-based visual saliency estimation," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1513–1525, Apr. 2012.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[3] P. Le Callet and E. Niebur, "Visual attention and applications in multimedia technologies," *Proc. IEEE*, vol. 101, no. 9, pp. 2058–2067, Sep. 2013.

[4] Y. Gao, M. Shi, D. Tao, and C Xu, "Database saliency for fast image retrieval," *IEEE Trans. Multimedia*, vol. 17, no. 3, pp. 359–369, Mar. 2015.

[5] X. Wei, Z. Tao, C. Zhang, and X. Cao, "Structured saliency fusion based on Dempster–Shafer theory," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1345–1349, Sep. 2015.

[6] M. Wang, R. Hong, X.-T. Yuan, S. Yan, and T.-S. Chua, "Movie2comics: Towards a lively video content presentation," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 858–870, Jun. 2012.

[7] S. Han and N. Vasconcelos, "Image compression using object-based regions of interest," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 3097–3100.

[8] R. Gallea, E. Ardizzone, and R. Pirrone, "Physical metaphor for streaming media retargeting," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 971–979, Jun. 2014.

[9] U. Rutishauser, D. Walther, C. Koch, and P. Perona, "Is bottom-up attention useful for object recognition?," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, Jun. 2004, pp. 37–44.

[10] J. Lei *et al.*, "Depth sensation enhancement for multiple virtual view rendering," *IEEE Trans. Multimedia*, vol. 17, no. 4, pp. 457–469, Apr. 2015.

[11] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[12] O. Le Meur, P. Le Callet, D. Barba, and D. Thoreau, "A coherent computational approach to model the bottom-up visual attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 802–817, May 2006.

[13] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc. Neural Inform. Process. Syst.*, Dec. 2006, pp. 545–552.

[14] W. Wang, Y. Wang, Q. Huang, and W. Gao, "Measuring visual saliency by site entropy rate," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2010, pp. 2368–2375.

[15] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 185–198, Jan. 2010.

[16] B. Schauerte and R. Stiefelhagen, "Quaternion-based spectral saliency detection for eye fixation prediction," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2012, pp. 116–129.

[17] J. Li, Y. Tian, and T. Huang, "Visual saliency with statistical priors," *Int. J. Comput. Vis.*, vol. 107, no. 3, pp. 239–253, May 2014.

[18] W. Einhäuser, M. Spain, and P. Perona, "Objects predict fixations better than early saliency," *J. Vis.*, vol. 8, no. 14, pp. 1–26, 2008.

[19] T. Liu, J. Sun, N. Zheng, X. Tang, and H. Y. Shum, "Learning to detect a salient object," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jul. 2007, pp. 1–8.

[20] H. Jiang *et al.*, "Salient object detection: A discriminative regional feature integration approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2013, pp. 2083–2090.

[21] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2009, pp. 1597–1604.

[22] E. Rahtu, J. Kannala, M. Salo, and J. Heikkila, "Segmenting salient objects from images and videos," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2010, pp. 366–379.

[23] V. Gopalakrishnan, Y. Hu, and D. Rajan, "Salient region detection by modeling distributions of color and orientation," *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 892–905, Aug. 2009.

[24] Z. Liu, Y. Xue, L. Shen, and Z. Zhang, "Nonparametric saliency detection using kernel density estimation," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 253–256.

[25] Z. Liu *et al.*, "Unsupervised salient object segmentation based on kernel density estimation and two-phase graph cut," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 1275–1289, Aug. 2012.

[26] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2010, pp. 2376–2383.

[27] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 409–416.

[28] V. Gopalakrishnan, Y. Hu, and D. Rajan, "Random walks on graphs for salient object detection in images," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3232–3242, Dec. 2010.

[29] Z. Liu, X. Zhang, S. Luo, and O. Le Meur, "Superpixel-based spatiotemporal saliency detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 9, pp. 1522–1540, Sep. 2014.

[30] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2012, pp. 733–740.

[31] G. Li and Y. Yu, "Visual saliency based on multiscale deep features", in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2015, pp. 1–9.

[32] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2007, pp. 1–8.

[33] J. Li, M. D. Levine, X. An, X. Xu, and H. He, "Visual saliency based on scale-space analysis in the frequency domain," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 996–1010, Apr. 2013.
[34] Y. Fang *et al.*, "Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 187–198, Feb. 2012.
[35] N. Imamoglu, W. Lin, and Y. Fang, "A saliency detection model using low-level features based on wavelet transform," *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 96–105, Jan. 2013.
[36] Y. Fang, W. Lin, Z. Chen, and C.-W. Lin, "Saliency detection in the compressed domain for adaptive image retargeting," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3888–3901, Sep. 2012.
[37] B. W. Tatler, "The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions," *J. Vis.*, vol. 7, no. 14, pp. 1–17, Nov. 2007.
[38] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
[39] V. Navalpakkam, C. Koch, A. Rangel, and P. Perona, "Optimal reward harvesting in complex perceptual environments," in *Proc. Nat. Acad. Sci.*, Mar. 2010, vol. 107, no. 11, pp. 5232–5237.
[40] R. Achanta *et al.*, "SLIC superpixels," EPFL, Lausanne, Switzerland, Tech. Rep. 149300, 2010.
[41] R. Achanta *et al.*, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[42] P. Ganesan, V. Rajini, and R. Rajkumar, "Segmentation and edge detection of color images using cielab color space and edge detectors," in *Proc. Int. Conf. Emerging Trends Robot. Commun. Technol.*, Dec. 2010, pp. 393–397.
[43] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2013, pp. 1155–1162.
[44] Y. Zhai and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues," in *Proc. ACM Int. Conf. Multimedia*, Oct. 2006, pp. 815–824.
[45] R. Margolin, A. Tal, and L. Zelnik-Manor, "What makes a patch distinct?," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2013, pp. 1139–1146.
[46] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2014, pp. 248–255.

Jianjun Lei (A'11-M'12) received the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications, Beijing, China, in 2007.

From August 2012 to August 2013, he was a Visiting Researcher with the Department of Electrical Engineering, University of Washington, Seattle, WA, USA. He is currently a Professor with the School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include 3D video processing, 3D display, and computer vision.

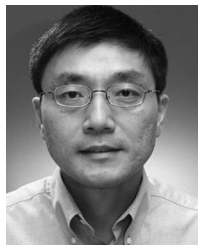
Bingren Wang received the B.S. degree in telecommunication engineering from Tianjin University, Tianjin, China, in 2012, where he is currently working toward the M.S. degree in electronic information engineering.

His research interests include 3D imaging and digital image processing.

Yuming Fang received the B.E. degree from Sichuan University, Chengdu, China, in 2006, the M.S. degree from the Beijing University of Technology, Beijing, China, in 2009, and the Ph.D. degree in computer engineering from the Nanyang Technological University, Singapore, in 2013.

He is currently an Associate Professor with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. He has authored or coauthored more than 70 academic papers. His research interests include visual attention modeling, visual quality assessment, image retargeting, and 3D image/video processing.

Prof. Fang is an Associate Editor of IEEE ACCESS. He has also served as a Session/Tracking Chair in many academic conferences.



Weisi Lin (S'91-M'92–SM'00–F'16) received the Ph.D. degree from King's College, London University, London, U.K., in 1992.

He previously served as the Laboratory Head of Visual Processing, Institute for Infocomm Research, Singapore. He is currently an Associate Professor with the School of Computer Engineering, Institute for Infocomm Research. He has authored or coauthored 160 journal papers, 230 conference papers, 2 books, and 9 book chapters, edited 3 books, and filed 7 patents. His research interests include perceptual modeling and evaluation of multimedia signals, image processing, and video compression.

Prof. Lin is a Chartered Engineer, a Fellow of IET, and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and the *Journal of Visual Communication and Image Representation*, and was an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and the IEEE SIGNAL PROCESSING LETTERS. He served as a Guest Editor of seven special issues in different scholarly journals. He has been a Technical Program Chair for IEEE ICME 2013, PCM 2012, QoMEX 2014, and VCIP 2017. He chaired the IEEE MMTC Special Interest Group on QoE (2012–2014). He has been a keynote/invited/panelist/tutorial speaker for more than 20 international conferences, as well as a Distinguished Lecturer of the IEEE Circuits and Systems Society (2016–2017) and the Asia-Pacific Signal and Information Processing Association (2012–2013).

Patrick Le Callet received the M.Sc. and Ph.D. degrees in image processing from École Polytechnique de l'Université de Nantes, Nantes, France, in 1993 and 2001, respectively.

He was an Assistant Professor from 1997 to 1999 and a full-time Lecturer from 1999 to 2003 with the Department of Electrical Engineering, Technical Institute, University of Nantes (IUT), Nantes, France. Since 2003, he has been a Teacher with the Electrical Engineering and Computer Science Department, Ecole Polytechnique de l'Université de Nantes. He is currently a Full Professor with the Computer Science Department, Ecole Polytechnique de l'Université de Nantes. Since 2006, he has been the Head of the Image and Video Communication Lab, CNRS IRCCyN, Nantes, France. He is also a Co-Chair within the Video Quality Expert Group, for the "Joint-Effort Group" and "3DTV" activities. He has authored or coauthored more than 200 publications and communications, and holds 13 international patents. His research interests include the application of human vision modeling in image and video processing, 3D image and video quality assessment, watermarking techniques, and visual attention modeling and applications.

Prof. Le Callet is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the Springer EURASIP *Journal on Image and Video Processing*, and SPIE *Electronic Imaging*.

Nam Ling (S'85–M'88–SM'99–F'08) received the B.Eng. degree from the National University of Singapore, Singapore, in 1981, and the M.S. and Ph.D. degrees from the University of Louisiana, Lafayette, LA, USA, in 1985 and 1989, respectively.

From 2002 to 2010, he was an Associate Dean of the School of Engineering, Santa Clara University, Santa Clara, CA, USA. He is currently the Sanfilippo Family Chair Professor and the Chair of the Department of Computer Engineering, Santa Clara University. He is also a Consulting Professor with the National University of Singapore; a Guest Professor with Tianjin University, Tianjin, China; a Guest Professor with Shanghai Jiao Tong University, Shanghai, China; a Cuiying Chair Professor with Lanzhou University, Gansu, China; and a Distinguished Professor with the Xi'an University of Posts and Telecommunications, Shaanxi, China. He has authored or coauthored over 180 publications and standard contributions, including 2 books in the fields of video coding and systolic arrays. He has filed or holds more than 15 U.S. patents.

Dr. Ling is a Fellow of IET. He was a Keynote Speaker for IEEE APCCAS, VCVP (twice), JCPC, IEEE ICAST, IEEE ICIEA, IET FC Umedia, IEEE Umedia, and a Workshop at XUPT, as well as a Distinguished Speaker for IEEE ICIEA. He has served as a General Chair/Co-Chair for IEEE Hot Chips, VCVP (twice), IEEE ICME, IEEE Umedia (thrice), and IEEE SiPS. He has also served as a Technical Program Co-Chair for IEEE ISCAS, APSIPA ASC, IEEE APCCAS, IEEE SiPS (twice), DCV, and IEEE VCIP. He was a Technical Committee Chair for IEEE CASCOM TC and IEEE TCMM, and has served as a Guest Editor or an Associate Editor for publications such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, the Springer *Journal of Signal Processing Systems*, and Springer *Multidimensional Systems and Signal Processing*. He was named an IEEE Distinguished Lecturer twice and was also an APSIPA Distinguished Lecturer. He was the recipient of the IEEE ICCE Best Paper Award (First Place) and six awards from Santa Clara University, four at the University level (Outstanding Achievement, Recent Achievement in Scholarship, President's Recognition, and Sustained Excellence in Scholarship) and two at the School/College level (Researcher of the Year and Teaching Excellence).

Chunping Hou received the M.Eng. and Ph.D. degrees in electronic engineering from Tianjin University, Tianjin, China, in 1986 and 1998, respectively.

Since 1986, she has been with in the School of Electronic and Information Engineering, Tianjin University, where she is currently a Full Professor and the Director of the Broadband Wireless Communications and 3D Imaging Institute. Her research interests include 3D image processing, 3D display, wireless communication, and the design and applications of communication systems.