ROBUST SALIENCY PROPAGATION BASED ON RANDOM WALKS

Chao Li[1,2], *Panwen Yang*[1], *Hao Sheng*[1]

[1]State Key Laboratory of Software Development Environment, School of Computer Science and Engineering, Beihang University, Beijing 100191, P.R.China
[2]Shenzhen Key Laboratory of Data Vitalization, Research Institute in Shenzhen, Beihang University, Shenzhen, P.R. China

ABSTRACT

A variety saliency detection methods have been based on the background prior knowledge. Nevertheless, some image patches connected to the image boundary are foreground noise. In this paper, we propose a saliency detection algorithm which propagates coarse saliency based on robust background prior set via random walks. First, we get the robust background prior set from the initial background set which is made up by all image patches connected to the image boundary, and compute a coarse saliency map through summation of global contrast between image patch in complement set of robust background prior set and image patch in robust background set. Second, we get foreground prior set by segmenting the coarse saliency via adaptive threshold, and propagate foreground prior set via random walks. Propagated foreground prior set is integrated by Bayesian inference model and resulting a saliency map. Third, we update saliency map by hypergraph and integrate current saliency and updated saliency by weighted mean summation. Finally, results from multi-scale saliency maps are integrated by pixel-wise weighted summation. Experimental result shows that our approach outperforms state-of-art approaches on two public available dataset.

***Index Terms*—** saliency detection, superpixels, robust background, random walks, Bayesian inference model,

1. INTRODUCTION

Saliency detection in an image is an useful technology which uses low-level feature to estimate the position, size and shape of object. Generic feature , distinct color, different texture and diverse shape etc, extracted from the region of salient object always originates from contrast to its neighbor region and background part. Saliency detection models intrinsically salient stimuli to estimate probability of the region of image is categorized as part of foreground. The primarily human attention can automatically judge the importance of different part of image, and conduct following processes on the salient part which is most importance in image. The saliency detection in image is significant importance to reduce the computation cost as a preprocess in the tasks of computer vision.

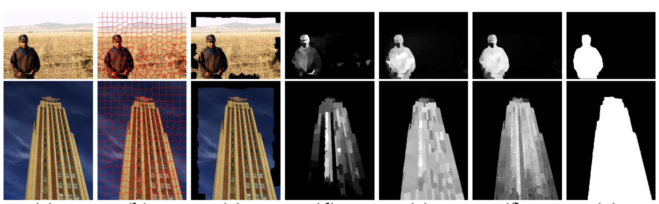
Since Itti et al. [1] introducing a computational model for visual attention based on integrated biological feature,saliency detection have had a rapidly increasing progress. In [2], Achanta et al. propose a conceptually simple approach by combining images band-pass filter responses from three $CIEL^*a^*b^*$ channels to model the saliency detection algorithm. In [3], Cheng et al. present an algorithm to exploit the pixel-wise saliency based on the contrast of color histogram (HC). Nevertheless, implementation of HC does not take the spatial information into consideration, the implementation of region contrast (RC) uses the spatial difference by the center point distance of image region to optimize the result of HC. Due to absence of high level feature and prior knowledge, previous saliency detection is difficult to deal with wide variation image. There are two means to model prior knowledge. First,the prior knowledge is based on foreground. In [4], Xie et al. model foreground prior by color boosting Harris corner detection [5], saliency map is integrated by Bayesian inference model.In [6],Liu et al. propose a saliency diffusion method by adaptive partial differential equation, the foreground prior is get by the Harris convex hull proposed by [4].Second, some saliency detection are based on the background prior. In [7], Wei et al. prove that image patch connected to the image boundary has a higher probability to be categorized as background. Furthermore, Wei et al. optimize the background prior knowledge by the property of boundary connectivity, an image patch is viewed as background only when the region is heavily connected to image boundary, in [8]. In [9], image patches connected to image boundary are used as a query vector to rank the probability of all image patches by manifold ranking. Based on absorbing Markov chain, the work of [10] models the saliency detection as excepted number of times transferring from transient state, superpixel disconnect to image boundary, to absorbing state, superpixel connected to image boundary.

In this paper, we propose a novel method to propagate saliency. All image patches connected to image boundary are collected as an initial set of background. Local contrast of the image patch in the initial background set is computed and grouped by the adaptive threshold [11]. B, robust background prior set, is made up by the image patch whose contrast value is less than the adaptive threshold. Image patch in complement set of robust background prior set is defined as U. We compute coarse saliency map via summing graph shortest path from image patch in U to image patch in robust background prior set B. To segment the coarse saliency map,we get an adaptive threshold [11] based on coarse saliency map. The image patch in U whose global contrast greater than the adaptive threshold [11] is used to set up foreground prior set. Random walks is used to propagate the foreground prior set and integrated the propagated saliency by Bayesian inference model. Then, we integrate current saliency map and updated saliency map by weighted mean summation. The final saliency map results from multi scale saliency integration by pixel-wise weighted summation.

2. OUR PROPOSED ALGORITHM

2.1. Robust Background Prior Set And Foreground Prior Set Estimation

In each scale of image segmentation, input image is segmented into N_i superpixels. An example of segmented result is shown in Fig. $1(b)$. There are some foreground noise touched to image boundary, as shown in Fig. $1(c)$. Using such a problematic background set lead negative effect to the following computation.



(a) (b) (c) (d) (e) (f) (g)

Fig. 1. (a) input image, (b) superpixel image with compactness is 20 and initial number of superpixel region is 250, (c) robust background prior set, (d) result of section 2.2, (e) result of section 2.3, (f) final result, (g) ground truth

We propose an effective method to remove foreground noise which touches to image boundary. We collect all superpixels connected to image boundary as an initial background set. Foreground noise tends to have distinctive contrast in the initial background set. As the contrast is based on adjacency relationship, we compute the contrast via local contrast.

$$b_i = \sum_{j}^{nei} \frac{d(c_i, c_j)}{1 + \alpha * d(s_i, s_j)} \tag{1}$$

i is the superpixel belongs to the initial background set. nei is an adjacent superpixel set consisted by neighbour vertices of superpixel v_i. At the same time, all element in nei is required to connect to image boundary. $d(c_i, c_j)$ is Euclidean distance of mean $CIEL^*a^*b^*$ color channel between superpixel i and superpixel j. $d(s_i, s_j)$ is Euclidean distance of spatial between center point of superpixel. α is used to adjust the importance between color contrast and spatial contrast. We remove superpixel whose local contrast is greater than adaptive threshold [11] from initial background set. The remain element in the initial background set are used to make up robust background prior set B. The result in Fig. $1(c)$ indicates that our method remove the foreground noise effectively.

Based on the set B, we build up a coarse saliency map via the global contrast. The salient region have an obvious difference compared with background. Nevertheless, summation of color contrast between different region lead to the cumulative error and can not encode the influence of the spatial.As all superpixels is segmented into two classes: robust background prior set, B, and the complement set, U. To encode color contrast and spatial information, we use graph shortest path to compute a coarse saliency map.

$$G_s = min_{p_1,p_2,\ldots,p_n} \sum \frac{d(c_{p_i}, c_{p_{i+1}})}{1 + \alpha * d(s_{p_i}, s_{p_{i+1}})} \tag{2}$$

p_i is adjacent to p_{i+1}. G_s is a matrix consist of graph shortest path to each vertices of the graph. The coarse saliency map is computed with following criteria: 1 the superpixel in robust background prior set is set zero, 2 the non-background superpixel is computed by summing the graph shortest path from the vertices in U to B.

$$S_i = \begin{cases} 0 & v_i \in B \\ \sum_{v_j \in B} G_{s_{i,j}} & v_i \notin B \end{cases} \tag{3}$$

We use adaptive threshold[11] to segment the non-zero value in coarse saliency map. The non-zero saliency value greater than threshold is marked as element of foreground prior set F.The vertices v_i in foreground prior set is assigned a value to measure the probability of v_i to be classified as a foreground.To get the probability of vertices in foreground prior set belongs to foreground, the saliency map of foreground prior set is normalized into [0, 1] as a prior probability set F_p, element f_{p_i} in F_p is used to measure the probability of v_i belongs to foreground set.

2.2. Coarse Saliency Map Propagated By Random walks

In [12], random walks is used to category the data under the condition of having initial labels. To propagate the coarse saliency map, we use robust background prior set B and foreground prior set F as initial condition to group the unmark data. To avoid negative propagation of the foreground prior, all vertices in the robust background prior set is labeled as a class to classify the unmark data, but we select only an element in foreground prior set as label in each iteration of classifying the unmark data.

$$S_{r_i} = -L_U^{-1} * B^T * \left[\begin{array}{cc} X_i & X_{B_a} \end{array}\right] \tag{4}$$

i is an vertices v_i in the foreground prior set $F.B_a$ is all vertices in the robust background prior set. L_U is part Laplacian matrix of unmark data. Weight of edge in Laplacian matrix,$w_{s_{i,j}}$,is defined by gaussian weighting function.

$$w_{s_{i,j}} = e^{-\frac{\|c_i, c_j\|^2}{\sigma^2}} \tag{5}$$

where c_i and c_j are the feature of the two vertices v_i and $v_j.\sigma$ is a parameter controls smoothness of gaussian weighting function.

S_{r_i} is condition probability of v_i is foreground. In Bayesian inference model, the posterior probability is computed by:

$$p(i|sal) = \frac{p(sal_i|i) * p(i)}{\sum_k^{n_p} p(sal_k|k) * p(k)} = \frac{p(sal_i|i) * p(i)}{p(sal)} \tag{6}$$

$p(i|sal)$ is posterior probability of vertices v_i under the condition of saliency map. $p(i)$ is the prior probability that vertices v_i is classified to be foreground part. $p(sal|i)$ is the likelihood of observation that under the condition vertices v_i is foreground. n_p is the number of potential foreground.

The saliency map propagated by random walks is integrated via Bayesian inference model. S_{r_i} is viewed as observation probability of the Bayesian inference model, and the prior probability is the foreground prior f_{p_i} . The probability of saliency map can be computed by:

$$S_b = \sum f_{p_i} * S_{r_i} \tag{7}$$

The result in Fig. $1(e)$. shows that our model based on random walks.

2.3. Saliency Updated By Weighted Mean

Random walks not only propagates the foreground but also the background noise in foreground prior set. To smooth the propagated background noise, we use normalized cut to update the propagated saliency map. Propagated saliency map is viewed as an eigenvector of the Laplacian matrix and the result of partition is the final saliency in each scale. In [13],Zhou et al. solve normalized cut by hypergraph theory. We use hypergraph theory to update saliency.

$$S_u = (I - 0.5 * D^{-\frac{1}{2}} * W * D^{-\frac{1}{2}}) * S_b \tag{8}$$

I is an identity matrix. D is the degrees matrix of superpixels graph $G = (V, E).W$ is weight matrix of graph $G = (V, E)$.Based on the property of laplacian matrix, the procedure of updating does not store current state and last state. In [14], the current saliency is stored by a coherence matrix. The coherence matrix is a diagonal matrix with diagonal element is determined by local difference. Nevertheless, the saliency propagated by random walks is determined by local difference. To decline the effect of local difference and store current state, we take the current state into consideration via identity matrix rather than coherence matrix. The identity matrix is not influenced by neighbour nodes and can store current state effectively. The equation of update can be defined by:

$$S_{t+1} = \delta * I * S_t + (1 - \delta)U_t \tag{9}$$

δ balances the importance between current stat and the state updated by normalized cut. In our experiment, we set $\delta = 0.9$ to adjust current state and updated state. Compared the result of Fig. $1(d)$ and Fig. $1(e)$, the weighted mean summation prohibits the background noise in foreground set effectively.

2.4. Integrate Multi Scale Saliency

To handle the difference of scale, we integrate final saliency map for each pixel via weighted summation.

$$S_f = \frac{\sum_n^N W_n * S_n}{\sum_n^N W_n} \tag{10}$$

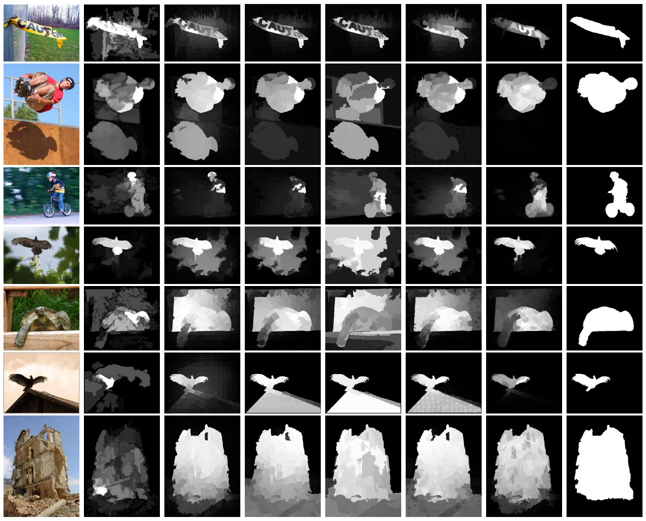
$$W_n = \frac{1}{\| pix_{n_c} - pix_{n_m} \|^2} \tag{11}$$

where N define the number of different scale. In our implement $N = 3$. S_n is the saliency map result from the weighted mean propagation at n-th scale. pix_{n_c} denotes the $CIEL^*a^*b^*$ channel of each pixel in the superpixel region and pix_{n_m} is the average $CIEL^*a^*b^*$ channel within the superpixel region.

3. EXPERIMENTAL RESULT

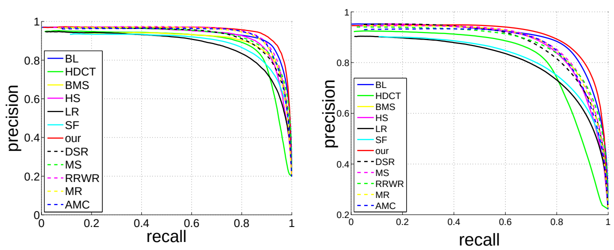
We exhaustively compare our approach with other state-of-the-art saliency detection methods on two public available dataset:ASD [2] and THUS [3]. For experimental comparison, we segment the input image into three scale with compactness 15, 20, 25 and initial number of superpixel 150, 250,350. The smoothing parameter σ used to conduct the gaussian weight value of random walks is assigned 25, the σ of normalized cut is 10. α = 3 to balance the importance of color contrast and spatial difference.Experimental results show that our method have a robust precision than other methods.

In Fig. 2, we provide some visual comparison of different saliency outputs result with previous saliency detection based on propagation: BMS[15], AMC[10], MR[9], HS[16], RRWR[17]. Our method not only removes foreground noise from image boundary set but also prohibits foreground error propagation from coarse saliency map rather than the other saliency detection methods.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 2. the visual comparison with state-of-art saliency detection algorithm. (a) Input image, (b) BMS[15], (c) AMC[10], (d) MR[9], (e) HS[16], (f) RRWR[17], (g) our, (h) ground truth.



(a) (b)

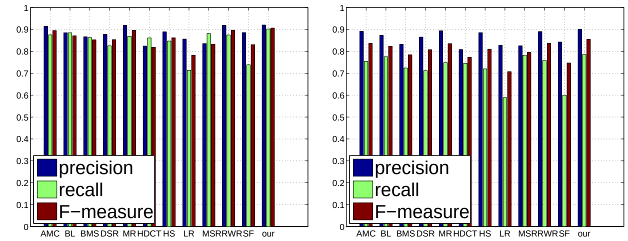
Fig. 3. precision-recall of (a)ASD and (b)THUS

We compare our model with eleven state-of-the-art saliency detection algorithms: LR[18], SF[19], BMS[15], AMC[10], MR[9], HS[16], DSR[20], HDCT[21], MS[22], BL[23], RRWR[17]. To save space, we only compare with top-performance methods for saliency detection in recent study. The result shows that our method outperforms the methods in precision recall curves, F-Measure and mean absolute error.

Similar to [2], we use a threshold T_f ranged in [0, 255] to segment saliency map. Subsequently, we use the binary to calculate precision and recall. When precision-recall pairs of all test images in data set are obtained, we generate an average precision recall pair. The P-R curves is shown in Fig. 3.

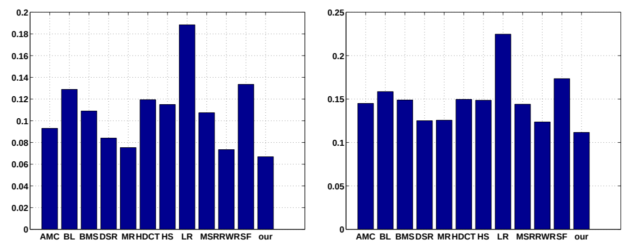
The saliency maps can be segmented by image dependent threshold proposed by [2]. Based on the threshold, we obtain a precision, recall and F-measure as in [2]. Averaged F-measure achieved by each saliency detection algorithm is listed in Fig. 4.

The P-R curves have a limitation that they only consider



(a) (b)

Fig. 4. precision, recall and F-measure of (a)ASD and (b)THUS



(a) (b)

Fig. 5. mean absolute error of (a)ASD and (b)THUS

whether the saliency region is higher than the background.The mean absolute error [8] is computed by average per-pixel difference between the binary ground truth and the saliency map. It is measured directly how closely the saliency map to ground truth.The result is shown in Fig. 5.

4. CONCLUSION

In this paper, we implement saliency detection method by propagated the coarse saliency via random walks. The saliency object has distinctive contrast to the background is used to remove the salient noise from the initial background set. The remained element of init background set is used to setup a robust background set. We use random walks to propagate the coarse foreground set based on the robust background set rather than the single-propagation in AMC[10], MR[9], RRWR[17]. At the same time, the prior foreground set can be viewed as prior probability to conduct a Bayesian inference model. Finally, the saliency map is smoothed by weighted mean summation.

Acknowledge . This study was partially supported by the National Natural Science Foundation of China (No. 61370122), the National High Technology Research and Development Program of China (No. 2013AA01A603) and the National Science & Technology Pillar Program (No.2015BAF14B01). Supported by the Programme of Introducing Talents of Discipline to Universities and the Open Fund of the State Key Laboratory of Software Development Environment under grant #SKLSDE-2015ZX-21.

5. REFERENCES

[1] Laurent Itti, Christof Koch, and Ernst Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on PAMI*, pp. 1254–1259, 1998.

[2] Ravi Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk, "Frequency-tuned salient region detection," in *CVPR*. IEEE, 2009, pp. 1597–1604.

[3] Ming Cheng, Niloy J Mitra, Xumin Huang, Philip HS Torr, and Song Hu, "Global contrast based salient region detection," *IEEE Transactions on PAMI*, pp. 569–582, 2015.

[4] Yulin Xie and Huchuan Lu, "Visual saliency detection based on bayesian model," in *ICIP*. IEEE, 2011, pp. 645–648.

[5] Joost Van de Weijer, Theo Gevers, and Andrew D Bagdanov, "Boosting color saliency in image feature detection," *IEEE Transactions on PAMI*, pp. 150–156, 2006.

[6] Risheng Liu, Junjie Cao, Zhouchen Lin, and Shiguang Shan, "Adaptive partial differential equation learning for visual saliency detection," in *CVPR*. IEEE, 2014, pp. 3866–3873.

[7] Yichen Wei, Fang Wen, Wangjiang Zhu, and Jian Sun, "Geodesic saliency using background priors," in *European Conference on Computer Vision*, pp. 29–42. Springer, 2012.

[8] Wangjiang Zhu, Shuang Liang, Yichen Wei, and Jian Sun, "Saliency optimization from robust background detection," in *CVPR*. IEEE, 2014, pp. 2814–2821.

[9] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang, "Saliency detection via graph-based manifold ranking," in *CVPR*. IEEE, 2013, pp. 3166–3173.

[10] Bowen Jiang, Lihe Zhang, Huchuan Lu, Chuan Yang, and Ming-Hsuan Yang, "Saliency detection via absorbing markov chain," in *ICCV*. IEEE, 2013, pp. 1665–1672.

[11] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *Automatica*, pp. 23–27, 1975.

[12] Leo Grady, "Random walks for image segmentation," *IEEE Transactions on PAMI*, vol. 28, no. 11, pp. 1768–1783, 2006.

[13] Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding," in *Advances in neural information processing systems*. 2006, pp. 1601–1608, Springer.

[14] Yao Qin, Huchuan Lu, Yiqun Xu, and He Wang, "Saliency detection via cellular automata," in *CVPR*, 2015, pp. 110–119.

[15] Jianming Zhang and Stan Sclaroff, "Saliency detection: A boolean map approach," in *ICCV*. IEEE, 2013, pp. 153–160.

[16] Qiong Yan, Li Xu, Jianping Shi, and Jiaya Jia, "Hierarchical saliency detection," in *CVPR*. IEEE, 2013, pp. 1155–1162.

[17] Changyang Li, Yuchen Yuan, Weidong Cai, Yong Xia, and David Dagan Feng, "Robust saliency detection via regularized random walks ranking," *CVPR*, pp. 2710–2717, 2015.

[18] Xiaohui Shen and Ying Wu, "A unified approach to salient object detection via low rank matrix recovery," in *CVPR*. IEEE, 2012, pp. 853–860.

[19] Federico Perazzi, Philipp Krähenbühl, Yael Pritch, and Alexander Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *CVPR*. IEEE, 2012, pp. 733–740.

[20] Xiaohui Li, Huchuan Lu, Lihe Zhang, Xiang Ruan, and Ming-Hsuan Yang, "Saliency detection via dense and sparse reconstruction," in *ICCV*. IEEE, 2013, pp. 2976–2983.

[21] Jiwhan Kim, Dongyoon Han, Yu-Wing Tai, and Junmo Kim, "Salient region detection via high-dimensional color transform," in *CVPR*. IEEE, 2014, pp. 883–890.

[22] Na Tong, Huchuan Lu, Lihe Zhang, and Xiang Ruan, "Saliency detection with multi-scale superpixels," *Signal Processing Letters*, pp. 1035–1039, 2014.

[23] Na Tong, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang, "Salient object detection via bootstrap learning," in *CVPR*, 2015, pp. 1884–1892.