Robust Saliency Detection via Regularized Random Walks Ranking

Changyang Li¹, Yuchen Yuan¹, Weidong Cai¹, Yong Xia², and David Dagan Feng¹
¹The University of Sydney, ²Northwestern Polytechnical University

Saliency detection refers to the exploration and localization of the region in a given image that attracts most human attention when presented [4]. In computer vision, saliency detection algorithms are usually categorized into bottom-up [1] and top-down [3] approaches, in which graph-based bottom-up approaches prevail the current research. However, many graph-based bottom-up saliency detection algorithms heavily depend on the boundary prior and the pre-processed superpixel segmentation, which in general is of low robustness, and leads to significant sacrifice of detail information from the original image.

In this paper, we propose a novel saliency detection method, which takes both region-based estimations and pixel-wise image details into account. Our innovations come in two aspects: on the one hand, we suggest the erroneous boundary removal, which filters out one of the four boundaries that most unlikely belonging to the background; on the other hand, we propose the regularized random walks ranking, which is independent of the superpixel segmentation, and can generate pixel-wised saliency maps that reflects full details of the image. The workflow of our algorithm is introduced as follows.

Firstly, we conduct the erroneous boundary removal. The input image is segmented into superpixels by the SLIC method [2], and all of the superpixels on each boundary are treated as a connected region. The original image is segmented into superpixels by the SLIC method [2], and all of the boundaries after the erroneous boundary removal will be used in the following saliency estimation. The saliency estimation result will be generated as $S(i), i = 1,...,n$, where $n$ is the total superpixel number.

Secondly, we conduct the saliency estimation with the graph-based manifold ranking [9], which consists of both background saliency estimation and foreground saliency estimation. However, only the three remaining boundaries after the erroneous boundary removal will be used in the background saliency estimation. The saliency estimation result will be generated as $S(i), i = 1,...,n$, where $n$ is the total superpixel number.

Finally, we conduct the proposed regularized random walks ranking, which is extended from the random walks model. The image is treated as a dataset $\mathcal{X} = \{x_1, ..., x_n\} \in \mathbb{R}^n$, where $n$ is the total pixel number. We first mark $s$ elements from $\mathcal{X}$ as the seed nodes. Without loss of generality, we assume that the first $s$ elements of $\mathcal{X}$ are the seeds, so $\mathcal{X} = \{x_{i_1}, x_{i_2}\}$, in which $x_{i_1}$ are the seed nodes and $x_{i_2}$ are the unseeded nodes. We define the weight matrix $W$, degree matrix $D$, and the Laplacian matrix $L$ similarly to [9]; then let $p^k = [(p^k_1, ..., p^k_n)]^\top$ denote the probability vector of $\mathcal{X}$ for label $k$, where $k=1,2$ stand for background/foreground, respectively. $p^k$ can thus be partitioned as $p^k = [(p^k_u, p^k_f)]^\top$, where $p^k_u$ is for the seed nodes, having fixed value as 1 for the corresponding label. The optimized $p^k$ is achieved by minimizing our modified Dirichlet integral,

$$
\text{Dir}[p^k] = \frac{1}{2} (p^k_u)^\top L (p^k_u) + \frac{1}{2} (p^k_f - \gamma (p^k_f - \gamma))^\top (p^k_f - \gamma) $$

$$
= \frac{1}{2} \left[ (p^k_u)^\top B p^k_u + \frac{1}{2} \left[ (p^k_u - \gamma)^\top (p^k_u - \gamma) \right] \right] + \frac{1}{2} \left[ (p^k_f - \gamma)^\top (p^k_f - \gamma) \right] \right)

$$

where $\gamma$ is a controlling parameter, and $\gamma$ is a pixel-wise indication vector inheriting the values of $S$. We define two thresholds $t_{\text{high}}$ and $t_{\text{low}}$,

$$
t_{\text{high}} = \frac{\text{mean}(S) + \text{max}(S)}{2}, \quad t_{\text{low}} = \text{mean}(S),
$$

to select seed pixels for $p^k_u$, with $Y > t_{\text{high}}$ as foreground seeds and $Y < t_{\text{low}}$ as background seeds. The optimized solution is obtained as,

$$
p^k = (p^k_u + \mu p^k_f)^\top (-B^T \mu p^k_f + \mu Y^k).
$$

$p^k_u$ and $p^k_f$ are then combined to form $p^k$. We set $k = 2$ to select the foreground possibility as the final foreground saliency output.

Figure 1 shows the effect of the proposed method in comparison with several state-of-the-art saliency detection methods. The details of our method are described in the full paper. The conclusion is that our method outperforms 12 state-of-the-art methods on two public datasets, in terms of both accuracy and robustness.


This is an extended abstract. The full paper is available at the CVF webpage (http://www.cv-foundation.org/openaccess/CVPR2015.py).