*ECAI 2016*
*G.A. Kaminka et al. (Eds.)*


*doi:10.3233/978-1-61499-672-9-1569*

# Salient Region Detection Based on the Global Contrast Combining Background Measure for Indoor Robots

**Na Li[1] and Zhenhua Wang[1] and Lining Sun[1] and Guodong Chen[1, 2]**

**Abstract.** In this paper, we propose a new method of salient region detection for indoor robots, which integrate the background distribution into the primary saliency. *Region roundness* is proposed to describe the compactness of a region to measure background distribution more robustly. In order to validate the proposed method, several influential ones are compared on the DSD dataset. The results demonstrate that the proposed approach outperforms existing methods and is useful for indoor robots.

## 1 INTRODUCTION

Vision system is the most important perception tool for the indoor robot. Saliency detection method is inspired by the primate-customized visual attention mechanism: select the most relevant information among the plethora of visual information [1]. Plenty of methods [2-4,6,9] perform well on the public datasets, but they cannot work well for indoor scene with complex backgrounds, several salient objects and illumination variations. Namely, the restrictions exist when these methods are applied on indoor robots.

In this paper, the work is focused on indoor robot application of detecting salient regions or objects. To address the problems mentioned above, a new salient region detection model is proposed, which comprises segmentation, primary saliency measure, background distribution measure and combination.

## 2 MOTHOD

The proposed method can be divided into four steps: segmentation, primary saliency measure, background distribution measure and combination.

For the first step, the graph-based RGB-D segmentation [7] (see also [5]), using depth and color feature, is introduced to keep the completeness of the salient objects.

For the second step, a color histogram (each channel in RGB color space is quantized to *N* bins) is built for each region $r_i$ and the representative colors $\boldsymbol{c_i} : \{c_i^1, c_i^2, ......, c_i^{n_i}\}$ are picked adaptively. Herein, $\boldsymbol{c_i}$ should satisfies:

$$\begin{cases} f(c_i^1) \ge f(c_i^2) \ge ...... \ge f(c_i^{n_i}) \\ \sum_{t=1}^{n_i-1} f(c_i^t) < \varphi, \sum_{t=1}^{n_i} f(c_i^t) \ge \varphi \end{cases} \tag{1}$$

In which $f(c_i^t)$ is occurrence frequency of color bin $c_i^t$ and $\varphi$ is the minimum of frequency cumsum for each region. We set N=8 and $\varphi = 75\%$.

The primary saliency will be measured next. Salient regions should be distinctive and high-contrast compared with other regions in the image. The primary saliency is evaluated by contrasting colors of a region with all others,

$$RS(r_i) = \sum_{i \ne j} \omega_j * D_c(i,j) \tag{2}$$

In which $D_c(i,j)$ is the color distance of region $r_i$ and $r_j$,

$$D_c(i,j) = \sum_{p=1}^{n_i} \sum_{q=1}^{n_j} f(c_i^p) * f(c_j^q) * \left\| c_i^p - c_j^q \right\| \tag{3}$$

Here we utilize occurrence frequency as the weight of the color distance to emphasize high-frequency ones. In equation (2), $\omega_j$ is the weight of region $r_j$,

$$\omega_j = AR(r_j) * e^{-\alpha * D_s(r_i, r_j)} \tag{4}$$

In which $AR(r_j)$ is the area ratio of region $r_j$ in the input image and $D_s(r_i, r_j)$ is the spatial distance between region $r_i$ and $r_j$. $\alpha$ is the scaling factor and set to 3.0 through the experiment.

For the third step, we discover that generally salient regions have the property of compact sizes while background ones distribute widely and near image boundaries. Based on the hypothesis, background distribution $BD(r_i)$ is introduced to reduce false negative and false positive from the previous step,

$$BD(r_i) = \omega_{BD}(r_i) * e^{-\beta * rd(r_i)} \tag{5}$$

In which $\beta$ controls the strength of $rd(r_i)$ and we use $\beta = 0.5$ in all experiments. $rd(r_i)$, called region roundness, is a new proposed measure to describe the compactness of a region and defined as:

$$rd(r_i) = rd(r_i) / L_c(r_i)^2 \tag{6}$$

In which $A(r_i)$ and $L_c(r_i)$ are area and contour length of region $r_i$ respectively. Region roundness is usually large for salient object regions and small for background regions, and Fig 1 (Each region roundness value is displayed on it) can support our inference.

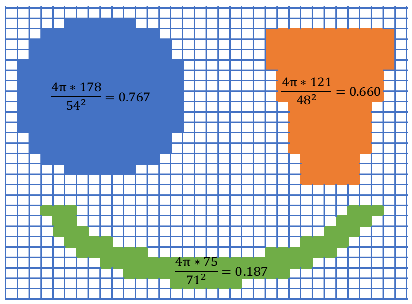


**Figure 1.** An illustrative example of region roundness.

[1] Robotics and Microsystem Center, Soochow University, Suzhou, China
[2] Corresponding author, email: guodongxyz@163.com

Spatial layout in images have the universality that background regions can be easily connected to image boundaries while foreground objects cannot [6]. Based on the hypothesis, we set a valid weight $\omega_{BD}(r_i)$ to influence background distribution measure,

$$\omega_{BD}(r_i) = 1 - e^{-\{BC(r_i)/\delta_{BC} + DC(r_i)/\delta_{DC}\}} \tag{7}$$

In which $DC(r_i)$ is the spatial distance between image center and region $r_i$, and $BC(r_i)$ is the boundary connectivity [6] of region $r_i$ to quantify how heavily a region is connected to image boundaries. $\delta_{BC}$ and $\delta_{DC}$ control $BC(r_i)$ and $DC(r_i)$, and are set as the maximum of $BC(r_i)$ and $DC(r_i)$ separately.

For the final step, to inhibit false positive of background regions and raise false negative of salient regions, the two normalized maps are combined in a way of an exponential function:

$$Sal(r_i) = RS(r_i) * e^{-\gamma * BD(r_i)} \tag{8}$$

In which $\gamma$ is the scaling factor and set to 2.0.

## 3 TESTS AND RESULTS

The proposed method is evaluated on DSD dataset [8] that consists of 80 RGB-D images obtained by a robot in a real-world indoor scene. Herein, we first give a visual comparison with five methods (GBVS [9], FT [2], RC [3], SF [4], RBD [6]) on this dataset. Corresponding saliency maps are shown in fig 2. Our method outperforms others visually in regard to the GT (ground truth).

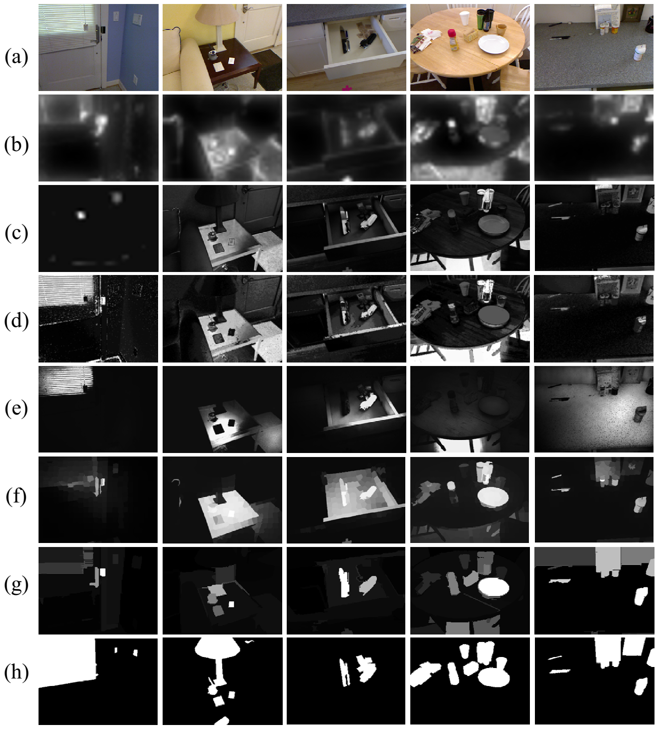


**Figure 2.** Visual comparison of saliency maps by different algorithms. (a) RGB image (b) GBVS (c) FT (d) RC (e) SF (f) RBD (g) Ours (h) GT

The PR (Precision and Recall) curves and MAE (Mean Absolute Error) are also introduced to quantifiably evaluate the proposed method. The PR curves are commonly used to reliably compare how well various saliency maps highlight salient regions in images. The resulting curves is shown in fig 3(a) and it clearly show that our method performs better than others. The MAE aims to measure how close a saliency map is to GT. Consistent with above results, the MAE results in fig 3(b) also support the excellence of our method.

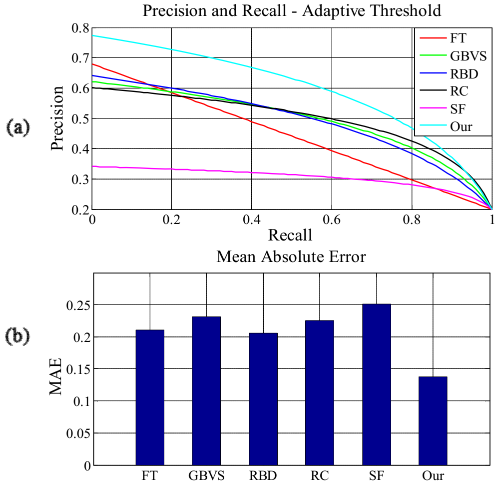


**Figure 3.** PR curves(a) and MAE(b) on DSD dataset for different methods.

## 4 CONCLUSIONS

This paper presents an effective method for indoor robots to detect salient regions or objects. However, parameter settings of segmentation weaken the application in indoor robots. We will optimize the segmentation more robustly as future work.

## ACKNOWLEDGEMENTS

The research is supported by NSFC (61305118 and U1509202) and JSNSF (BK20130329).

## REFERENCES

[1] M. Begum and F. Karray, 'Visual Attention for Robotic Cognition: A Survey', *IEEE Transactions on Autonomous Mental Development*, **3,** 92-105, (2011).

[2] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, 'Frequency-tuned Salient Region Detection', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1597-1604, 2009.

[3] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, 'Global Contrast Based Salient Region Detection', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 409-416, 2011.

[4] F. Perazzi, P. Krähenbühl, Y. Pritch, and A. Hornung, 'Saliency Filters: Contrast Based Filtering for Salient Region Detection', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 733-740, 2012.

[5] L. Jiang, A. Koch, and A. Zell, 'Salient Regions Detection for Indoor Robots using RGB-D Data', *IEEE International Conference on Robotics and Automation (ICRA)*, 1323-1328, 2015.

[6] W. Zhu, S.Liang, Y. Wei, and J. Sun, 'Saliency Optimization from Robust Background Detection', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2814-2821, 2014.

[7] P.F. Felzenszwalb, and D.P. Huttenlocher, 'Efficient Graph-Based Image Segmentation', *International Journal of Computer Vision*, **59**, 167-181, (2004).

[8] A. Ciptadi, T. Hermans, J.M. Rehg, 'An In Depth View of Saliency', *In Proceedings of the British Machine Vision Conference (BMVC)*, 9-13, 2013.

[9] J. Harel, C. Koch, and P. Perona. 'Graph-Based Visual Saliency', *Advances in neural information processing systems*, 545-552, 2006.