

## Research Article

# Obstacles Regions 3D-Perception Method for Mobile Robots Based on Visual Saliency

Tao Xu,<sup>1,2,3,4</sup> Songmin Jia,<sup>1,3,4</sup> Zhengyin Dong,<sup>1,3,4</sup> and Xiuzhi Li<sup>1,3,4</sup>

<sup>1</sup>College of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100124, China

<sup>2</sup>School of Mechanical and Electrical Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China

<sup>3</sup>Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

<sup>4</sup>Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124, China

Correspondence should be addressed to Tao Xu; [wlem1981@emails.bjut.edu.cn](mailto:wlem1981@emails.bjut.edu.cn)

Received 19 August 2015; Revised 19 November 2015; Accepted 19 November 2015

Academic Editor: Keigo Watanabe

Copyright © 2015 Tao Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel mobile robots 3D-perception obstacle regions method in indoor environment based on Improved Salient Region Extraction (ISRE) is proposed. This model acquires the original image by the Kinect sensor and then gains Original Saliency Map (OSM) and Intensity Feature Map (IFM) from the original image by the saliency filtering algorithm. The IFM was used as the input neuron of PCNN. In order to make the ignition range more exact, PCNN ignition pulse input was further improved as follows: point multiplication algorithm was taken between PCNN internal neuron and binarization saliency image of OSM; then we determined the final ignition pulse input. The saliency binarization region abstraction was fulfilled by improved PCNN multiple iterations finally. Finally, the binarization area was mapped to the depth map obtained by Kinect sensor, and mobile robot can achieve the obstacle localization function. The method was conducted on a mobile robot (Pioneer3-DX). The experimental results demonstrated the feasibility and effectiveness of the proposed algorithm.

## 1. Introduction

In recent years, the robots have already massively been applied in many fields. In particular, mobile robots have developed rapidly to help human complete kinds of difficult tasks. For mobile robots, to achieve efficient processing of visual information about the environment is obviously important, so the method which a vision-based obstacle avoidance system needs to recognize obstacles and describe them mathematically to give out the obstacle avoidance strategy has been a hot research focus.

Currently, the obstacle recognition and segmentation methods based on visual information mainly include four types [1]. Ohya et al. [2] had presented a model-based approach used for locating obstacles; obstacles are detected by computing the difference between the edges estimated from the 3D environment model and the edges detected from the actual camera image. Nevertheless, their results failed to detect the brown box in 50% of the cases. Aiming to solve these problems, we present an automatic obstacles

regions extraction algorithm, which is based on the saliency of visual attention. To mimic the astonishing capability of human visual system, visual saliency algorithm can easily detect salient visual subsets from images and videos. Camus et al. [3] developed a model that uses only real-time image flow divergence to avoid obstacles while driving toward a specified goal direction in a lab containing office furniture and robot and computing equipment. The major contribution of this work is the demonstration of a simple, robust, minimal system that uses flow-derived measures to control steering and speed to avoid collision in real time for extended periods. But this approach only uses observable image information represented in the 2D image sequence. It cannot convey adequate information on real users' environments needs. Compared with the existing methods, our method can achieve obstacles regions 3D-perception for mobile robot by effectively taking advantage of the Kinect sensor system. Hong et al. [4] described the local method which is based on feature correlation. According to certain criteria, "characteristic points" are selected from each view and paired

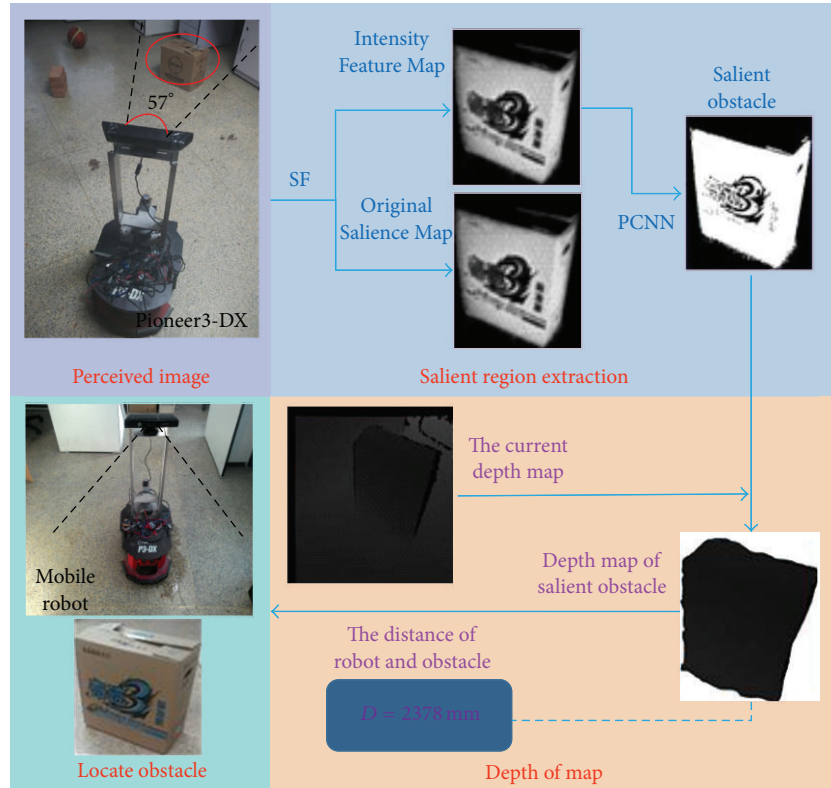


FIGURE 1: Frame diagram of our model's application.

with characteristic points in the other view by computing a measure of similarity between the two windows around the characteristic points. Bing et al. [5] built a model that the behavior state can generate according to the optical flow field changes of ground while the robot is moving. But for the noises, shadows and occlusion existing, the calculated optical flow is not precise and reliable enough, while the optical flow field computing is also time-consuming.

Considering the movements of robot, the relative position of obstacles area in robot's view would be changed; therefore, obstacles area extraction can be transformed into a motion region extraction problem. Based on this, in this paper, we present a new obstacle regions localization method in indoor environment, which is applied on mobile robot. To ensure the mobile robot can pass through the suitable path, the robot needs to detect and locate obstacles in the path. Firstly, scene images are generated by the Kinect sensor. Then, the mobile robot system gains OSM and IFM of the original image by the saliency filtering algorithm. Referring to the improved PCNN theory, we point that multiplication algorithm was taken between PCNN internal neuron and binarization saliency image of OSM; then we determine the final ignition pulse input. Finally, the saliency binarization region abstraction was fulfilled by improved PCNN multiple iterations. On this basis, the binarization area is mapped to the depth map obtained by Kinect sensor and the mobile robot can locate the obstacle [6].

The remainder of this paper is organized as follows. In Section 2, the overall design is addressed. Section 3

introduces ISRE method of SF algorithm with PCNN algorithm. Section 4 describes the advantage of ISRE algorithm on mobile robot followed by experiment results and analysis. Conclusions and future work are discussed in Section 5.

## 2. Overview of Our Model

To achieve obstacle information well, the new obstacle regions localization method is applied on mobile robot for its advantages such as recognition and location of obstacle. The detail of application is shown in Figure 1.

Broadly speaking, to accomplish the location task, this proposed method can be divided into three parts [7]. One is perceived image. The purpose of this part is to acquire scene images by the Kinect sensor. Another one is salient region extraction. This part aims to gain OSM and IFM of the original image by SF algorithm firstly, then IFM is being as the input neuron of PCNN, and finally saliency binarization region abstraction was fulfilled by improved PCNN. Last one is acquiring depth information and locating obstacle. For this part, the saliency binarization area is mapped to the depth map obtained by Kinect sensor; we can obtain depth map of the obstacle and locate the obstacle.

## 3. The Architecture of Algorithm

In order to locate obstacle for mobile robot, firstly the mobile robot needs to recognize the obstacle. In this section, we use the ISRE method combining SF algorithm [8] with PCNN

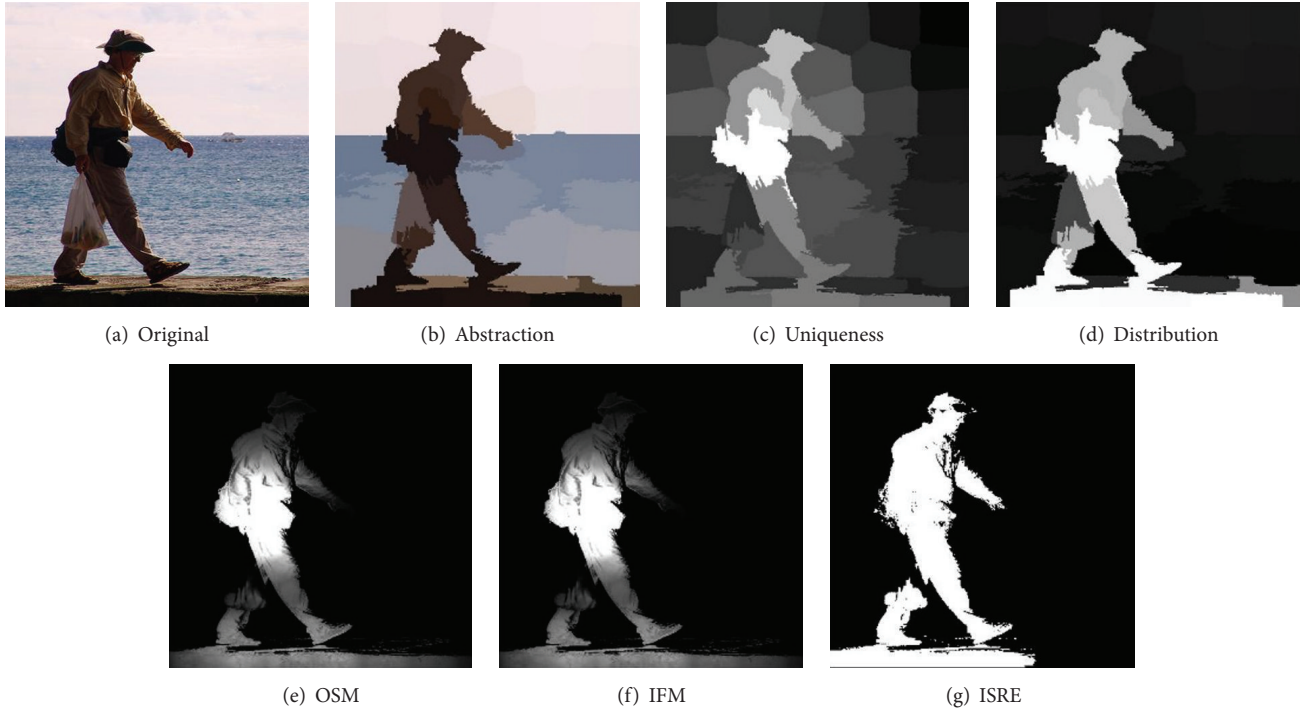


FIGURE 2: Effect map of units in ISRE model.

algorithm. The method extracts original saliency image and intensity feature image based on SF algorithm [8]; all pixels of intensity feature image are regarded as the input neurons of PCNN. In order to achieve the more precise sparking range, the ISRE algorithm improved the inputs of PCNN sparking pulse to complete the saliency binarization region abstraction by many times iterations. Then, according to the depth map of obstacle, the mobile robot can obtain the distance information to locate the obstacle.

**3.1. The ISRE Method.** As shown above, ISRE algorithm is a mixed model which simulates human eye's biological process. The ISRE algorithm has main 6 steps.

*Step 1.* Gain OSM and IFM through image's rough segmentation.

*Step 2.* Every pixel of IFM is regarded as the PCNN's input neuron.

*Step 3.* Every external stimulation forms local stimulation through connecting synaptic weight.

*Step 4.* Gain PCNN's internal items through connecting modulation.

*Step 5.* Improve PCNN's ignition pulse.

*Step 6.* Generate the binary saliency map through many iterations.

The process is shown in Figure 2. Figure 2(a) is the original image. The original image is first abstracted into perceptually homogeneous elements, as shown in Figure 2(b).

Each element is represented by the mean color of the pixels belonging to it. Then, two contrast measures per element based on the uniqueness and spatial distribution of elements were defined. The first contrast measure implement is a commonly employed assumption that image regions, which stand out from other regions in certain aspects, catch our attention and hence should be labeled more salient, as shown in Figure 2(c). The second contrast measure implement renders unique elements more salient when they are grouped in a particular image region rather than evenly distributed over the whole image, as shown in Figure 2(d). Figure 2(e) is the Original Saliency Map (OSM) and Figure 2(f) is the Intensity Feature Map (IFM). Figure 2(g) is the final saliency map with our ISRE method.

Step 1 will be as described in Section 3.2. Section 3.3 describes the concrete process of Steps 2 and 3. Step 4 will be discussed in Section 3.4, and Steps 5 and 6 will be presented in Section 3.5.

**3.2. Obstacle Saliency Region Extraction Unit.** Saliency map of SF algorithm [8] does not overcome the faults which the contrast difference model brings based on single global vision. In other words, when the high intensity region is in the background, it will extract the background as the saliency region, as shown in Figure 3(b). Referring to the question, ISRE model utilizes the segmentation idea that is from rough to fine. Firstly, we conduct rough segmentation with SF algorithm, then fine segmentation with the improved PCNN algorithm for the image, and complete the saliency binarization region abstraction finally, as shown in Figure 3(c).

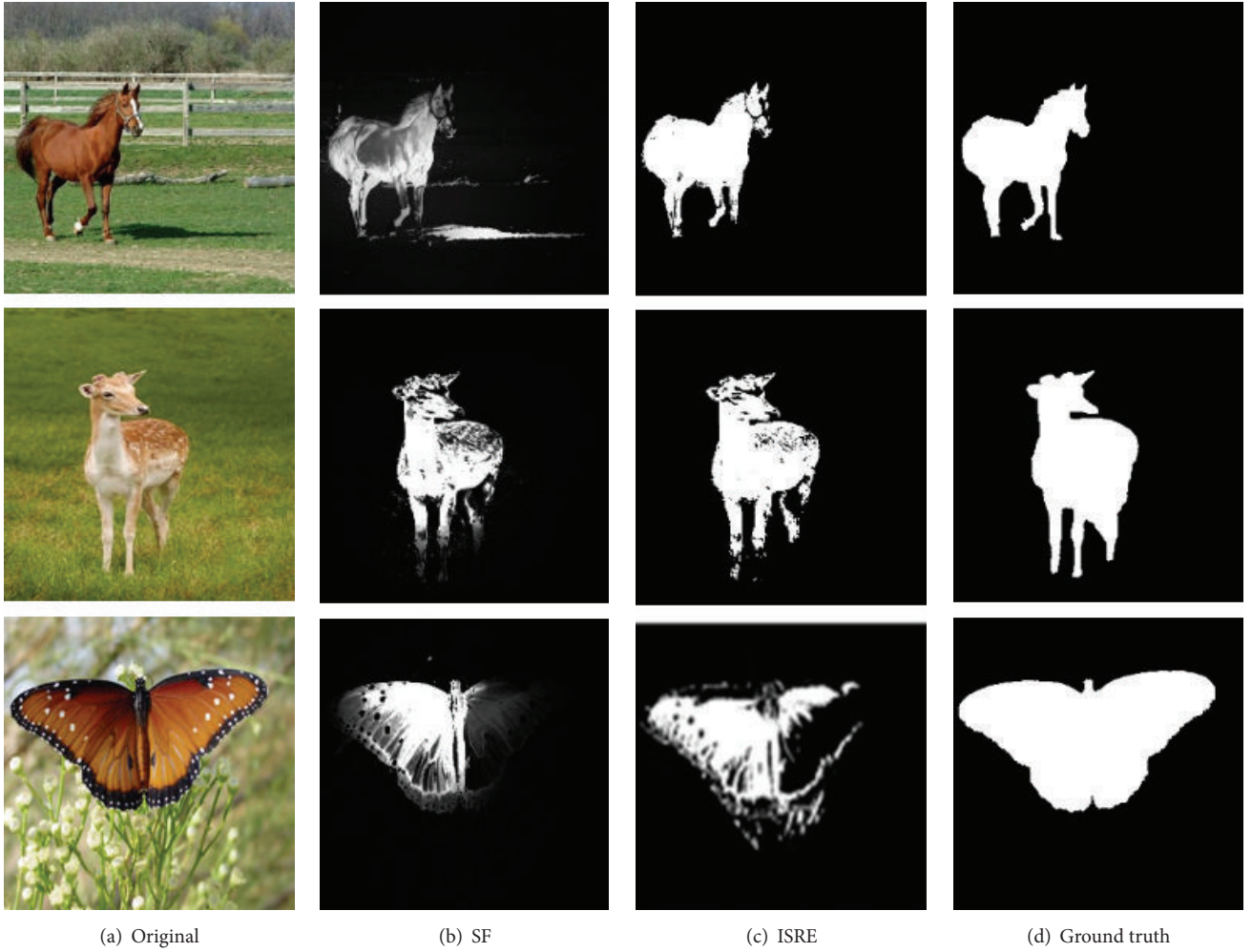


FIGURE 3: Effect contrast map of ISRE and SF.

**3.3. PCNN Input Unit.** The original image obtains OSM and IFM through SF algorithm. Its specific process is divided into four steps.

*Step 1.* Conduct super pixels segmentation for the input original image.

*Step 2.* Extract each element uniqueness of super pixels.

*Step 3.* Measure the distribution of the whole image; that is to say, refer to the certain region elements and highlight their saliency by rendering depth.

*Step 4.* Fuse Step 2 with Step 3 and determine OSM and IFM.

IFM's element  $I_{ij}(x, y)$  is regarded as PCNN's stimulation input  $F_{ij}$  [9]; meanwhile, synaptic connection weights  $W$  and a local stimulation  $L_{ij}$  [9] were formed by combining  $I_{ij}(x, y)$  with 8 neighborhoods of  $3 \times 3$  region. The value of  $W$  is determined by the distance from center pixel to its boundary pixel. The value of  $W$  is shown by

$$W = [0.707 \ 1 \ 0.707; 1 \ 0 \ 1; 0.707 \ 1 \ 0.707]. \quad (1)$$

**3.4. Modulation Unit.** The modulation unit's main function is coupling modulation of outside stimulation main input  $F_{ij}$  and local stimulation input connection  $L_{ij}$  as shown in [9]

$$U_{ij}[n] = F_{ij}[n] \{1 + \beta L_{ij}[n]\}, \quad (2)$$

where  $U_{ij}$  is internal items of neuron and  $\beta$  is the connecting strength coefficient among synapses. The value is higher; 8 neighborhoods' neurons have more effect on central neurons. The value of  $\beta$  is 0.4 in the paper.

**3.5. Pulse Generator Unit.** An important characteristic of the human biological visual process is that visual attention changes as the change of scene. In order to simulate the process accurately, the ISRE algorithm improved traditional PCNN ignition pulse unit in the paper. The specific steps can be described as follows.

*Step 1.* Take OSM's pixels which are 75% maximum nonzero gray value, and set their value as 1 (white), so as to define a maximum range of saliency region. The others' values are set

as 0 (black) to define certain background region. That is to say, we have achieved OSM's binary image OSM\_C.

*Step 2.* The neuron's internal items  $U_{ij}$  and OSM\_C conduct dot product to determine the final ignition pulse input as shown in

$$U_{kl} = U_{ij} \cdot^* \text{OSM\_C}. \quad (3)$$

*Step 3.* The  $U_{kl}$  determined neuron's dynamic threshold  $\Theta_{ij}$ . This process is presented as follows:

$$\Theta_{ij}[n] = \max[U_{kl}(x, y)] e^{-\alpha}, \quad (4)$$

where  $\alpha$  [9] is index attenuation coefficient of traditional PCNN's dynamic threshold.

The value of  $\alpha$  is 0.3 in the paper.

*Step 4.* Compare maximum of  $U_{kl}$  with dynamic threshold  $\Theta_{ij}$  to determine all firing neurons' value range and generate timing pulse sequence  $Y_{ij}$ , as shown in

$$Y_{ij}[n] = \begin{cases} 1, & U_{kl}[n] > \Theta_{ij}[n], \\ 0, & \text{Otherwise.} \end{cases} \quad (5)$$

Through many iterations, we finally complete the extraction of the binary saliency map. The accuracy of the ISRE is related to the number of iterations; the more the number of iterations, the higher the accuracy. But increasing the number of iterations will take a lot of time. After repeated tests, we choose three iterations in this paper. The number of iterations in ISRE model has the effect of inhibiting the human visual nervous system; that is to say, it could effectively suppress the noise from significant target zone through isolated neurons or tiny area of ignition.

**3.6. Obstacle Regions 3D-Perception.** With the advent of new sensing technology, 2D color image (RGB) along with the depth map (D) can now be obtained using low cost web camera style sensors such as Microsoft's Kinect [10]. The depth map provides per pixel depth information obtained using an infrared laser projector combined with a camera. While it does not provide a true 3D mesh, an RGB-D image provides more information about the captured object than a traditional 2D color image. RGB-D images have been applied to a variety of tasks such as general object recognition, surface modeling and tracking, modeling indoor environments, face recognition, discovering objects in scenes, and robotic vision [11–13]. Since the nature of an RGB-D image obtained from Kinect and a regular 3D map is fundamentally different, existing obstacle localization based on visual saliency techniques may not directly be applicable on these images. Each pixel in Kinect's depth map has a value indicating the relative distance of that pixel from the sensor at the time of image capture. It is important to utilize both RGB and depth data for feature extraction and classification.

In order to locate the obstacles, we obtained the depth map of obstacle through the binarization area that is mapped

to the depth map collected by Kinect sensor and then calculated the distance between obstacle and robot through

$$D = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M x_{ij}, \quad (6)$$

where  $D$  stands for the distance between robot and obstacle,  $x_{ij}$  means the value of the depth image, and  $M$  and  $N$  stand for the depth image's row and col.

## 4. Experimental Results

**4.1. Standard Database Results.** To verify that the ISRE algorithm is effective in extracting the image significantly, do contrast experiment in two aspects: visual effect and objective quantitative data comparison, based on the database which is about the 1000 original images and the significant true value offered by [14, 15] and the database about the seven existing salience extraction algorithms. Visual contrast experiment result is shown in Figure 4.

As can be seen from the figure, the second column IT algorithm [16] failed to fully highlight the significant area. The third column SR algorithm [17] only focuses on the extraction of significant edge, but lost inside information. The fourth column GB algorithm [18] and the fifth column MZ algorithm [19] are blurred for the extraction effect. The sixth column FT algorithm [15] and the seventh column AC algorithm [20] overall performance is better, but the effect is not good when significant background is too strong. For example, the grid of image in the third line and the bamboo of image in the fourth line are treated as a significant area and are extracted out. The eighth column SF algorithm [8] is the basis of ISRE model, and it extracts the significant area more accurately in various types of scenarios, while compared with the real value the extraction effect still needed to be improved when the significant region edge and the brightness have large differences. The ninth column is the final ignition pulse input result according to the algorithm mentioned in this paper; the overall effect has improved significantly on the basis of the SF algorithm; the extraction about the leaf areas of the image in the fourth row and the geese wings of the image in the second row has significantly improved and become closer to the value in tenth row (Ground Truth, GT).

To further verify the proposed algorithm more objectively, we need to consult [15] and design three experimental programs to quantitative calculation of IT [16], SR [17], GB [18], MZ [19], FT [15], CA [1], AC [20], SF [8], and ISRE algorithms to obtain objective comparative experimental data. All of the three protocols given by [12] make a comparison of the number of pixels in significant area extracted by the significant algorithm and the number of pixels in true area and then calculate their recall, precision, and  $F$ -Measure values which are shown in (7). The difference is that experiment 1 used custom threshold to divide the binary saliency map and every database's image result is displayed by curve graph, while experiment 2 used adaptive threshold to divide the binary saliency map and every database's image result is displayed by histograms graph. The result was shown in Figure 5.

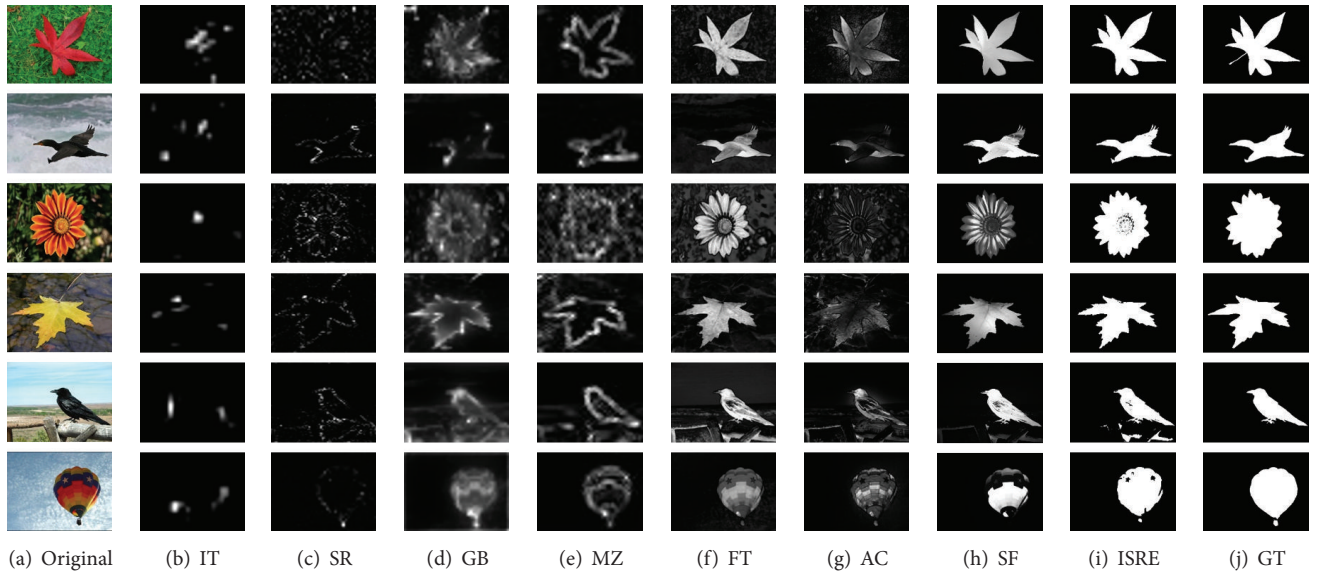


FIGURE 4: Contrast experiment among the seven existing algorithms, the ground truth, and ISRE.

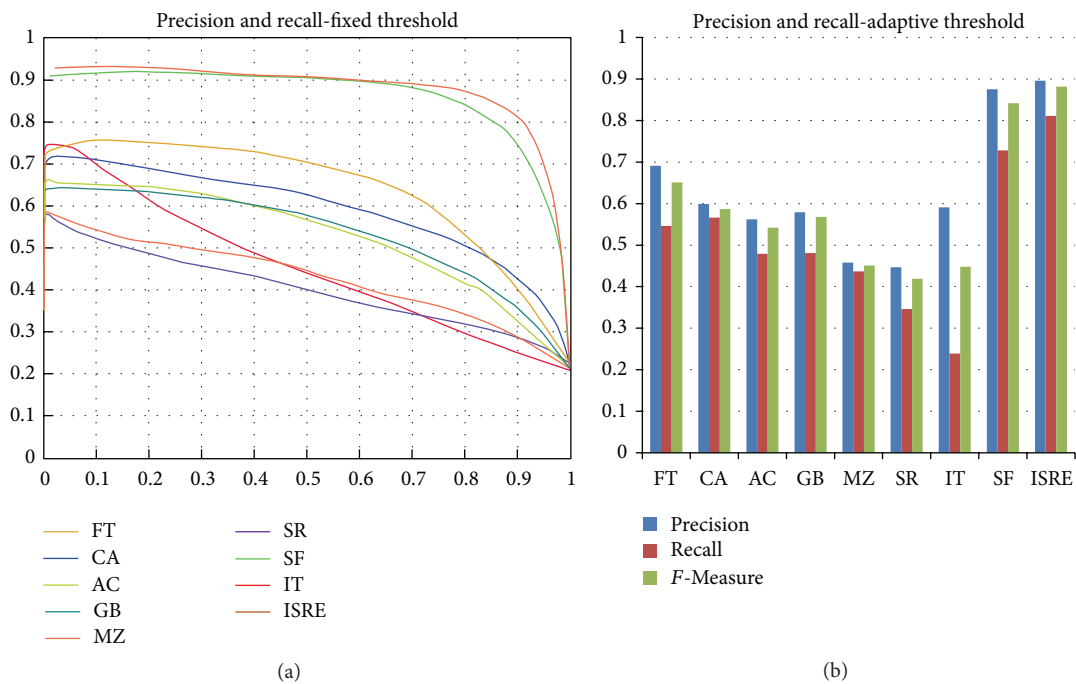


FIGURE 5: Comparison results of recall, precision, and  $F$ -Measure values.

DSR (Detected Salient Regions) are significant regions calculated by column. Refer to the standard mentioned in [21]; that is to say, to avoid having too high value of recall which maybe resulted by extracting much larger significant area, we should set  $\beta^2 = 0.3$  to enhance the proportion of precision.

It can be seen from Figure 5(a) that the extraction effects of each image in database based on the ISRE column mentioned in this paper were improved or unchanged compared to those based on SF algorithm. According to Figure 5(b), the precision ratio, recall ratio, and  $F$ -Measure of the ISRE

algorithm can reach 0.891, 0.808, and 0.870, which are obviously higher than the other eight columns:

$$\text{Recall} = \frac{\text{DSR} \cap \text{GT}}{\text{GT}},$$

$$\text{Precision} = \frac{\text{DSR} \cap \text{GT}}{\text{DSR}}, \quad (7)$$

$$F = \frac{(1 + \beta^2) \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}}.$$

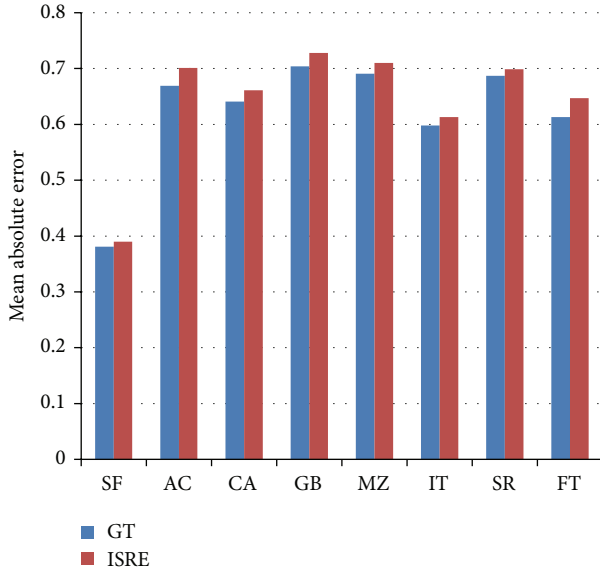


FIGURE 6: Comparison of mean absolute error relative to ISRE and GT.

**4.2. Mean Absolute Error.** In order to demonstrate that, the extraction results of ISRE algorithm are more approximate to the binary ground truth GT. For a more balanced comparison that takes these effects into account, we therefore also evaluate the mean absolute error (MAE) between the continuous saliency map  $S$  (prior to thresholding) and the binary ground truth GT. The mean absolute error is defined as

$$MAE1 = \frac{1}{a \times b} \sum_{i=1}^a \sum_{j=1}^b |F(i, j) - GT(i, j)|, \quad (8)$$

$$MAE2 = \frac{1}{a \times b} \sum_{i=1}^a \sum_{j=1}^b |F(i, j) - ISRE(i, j)|,$$

where  $a$  and  $b$  are the width and the height of the respective saliency map and ground truth image. Figure 6 shows that the same method MAE measure for ISRE and GT is very approximate. The IT algorithm MAE measure for ISRE is 0.610 and for GT it is 0.599. The MAE is almost equal. That demonstrates that our extraction is close to GT, as shown in Figure 6.

### 4.3. Actual Scene Results

**4.3.1. System Hardware Platform.** The platform used for environment exploration is an American Mobile Robots Inc.3 Pioneer3-DX embedded with a Kinect, illustrated in Figure 7.

The robot weighs over 9 kg, and the body of the robot is 44 cm × 38 cm × 22 cm. It possesses two driven wheels and one follower wheel, and the driven wheels' radius reaches 4.5 cm. The Kinect is a new and widely available device for the Xbox 360. The interest for Kinect is increasing in computer vision due to its advantages of providing 3D information of the environment and low cost. The device contains an RGB camera, a multiarray microphone, and a depth sensor. Using these sensors, Kinect executes depth measurement by light

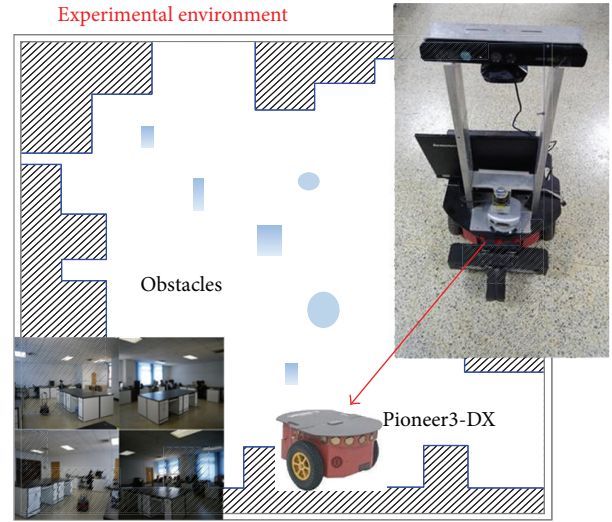


FIGURE 7: System hardware platform for environment exploration.

coding and captures full body 3D motion. Three software systems need to be installed to drive the Kinect, which are “OpenNI”, “SensorKinect”, and “NITE”. The version number of the “OpenNI” which we used is 1.5.2.23.

**4.3.2. Obstacles Regions 3D-Perception.** To further verify the effectiveness of the ISRE algorithm on mobile robot, we use this algorithm to extract significant area of real scene image in our laboratory, and extracted object is set to three; there are basketball, bricks, and carton; the experiment result was shown in Figure 8.

When the mobile is located in position A, the bricks are recognized by ISRE, and the distance between brick and robot can be calculated by (6); it is 1436 mm. We mark this position of the robot and measure the true distance with a ruler. The true distance is defined as  $d$ . The value is 1458 mm. In the same way, when the robot is located in positions B, C, D, and E it can recognize the bricks, basketball, and carton, and the distances, respectively, are 1273 mm, 1108 mm, 1704 mm, and 2378 mm, while the true distances, respectively, are 1301 mm, 1084 mm, 1741 mm, and 2359 mm. The relative error does not exceed 5%. We also can see from Figure 8 that this algorithm simulates human vision; it is possible to achieve identifying the target object accurately, detecting ground obstacle, and locating the obstacle functions. When the Pioneer3-DX moves, the obstacle distance information is gained through the depth of obstacle image meanwhile. While the extraction may be quite different in different scenes, there is a problem that some significant information might be lost.

**4.3.3. Contrast Experiment.** The obstacles regions 3D-perception strategy based on ISRE method is used to compare with SF method in the same scenario of our laboratory. The experiment results are demonstrated in Figure 9.

In this experiment, the robot starts at the red node position shown in the right half of Figure 9. Many obstacles

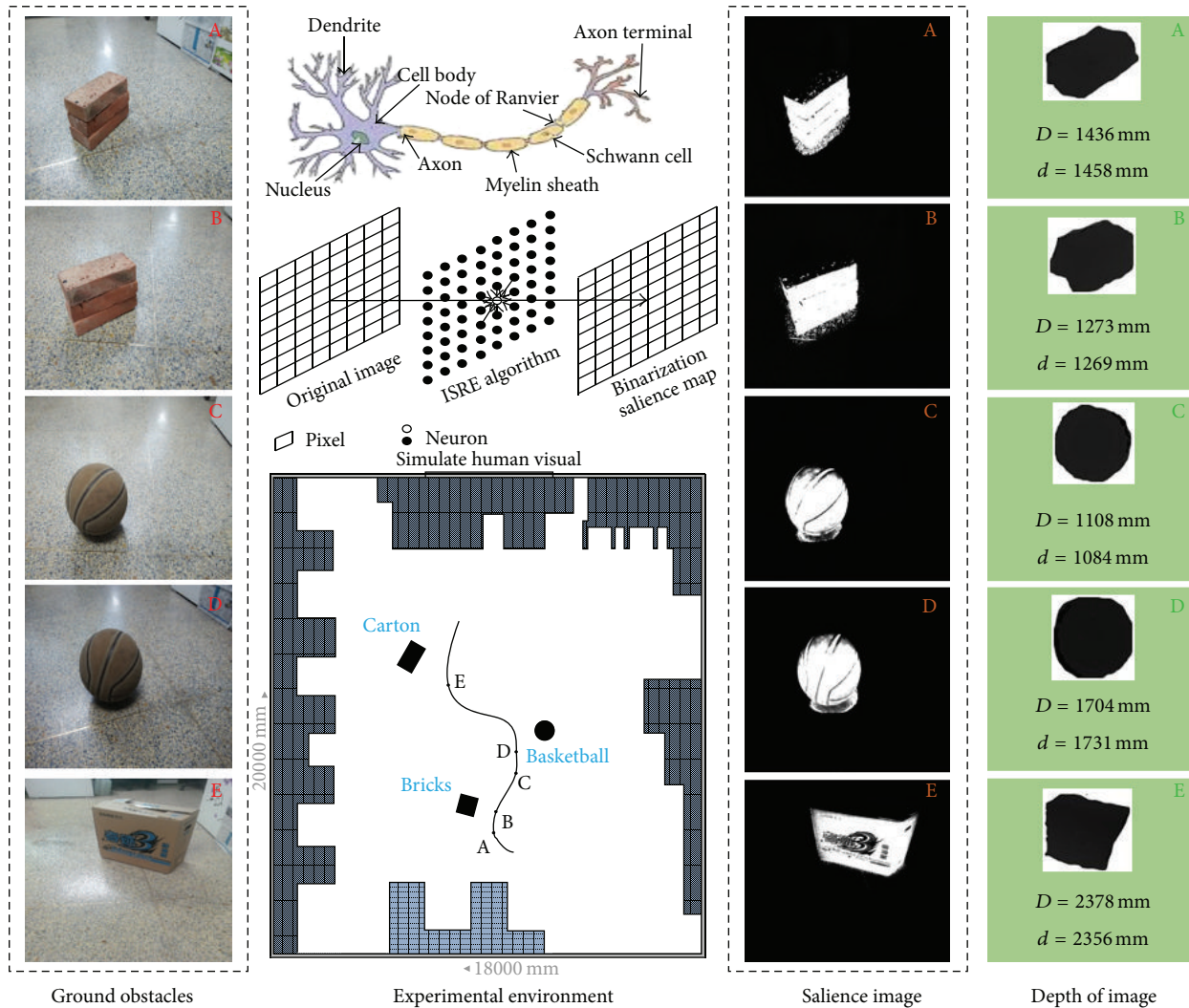


FIGURE 8: Experiment results of real world images with our algorithm.

were put carelessly in our laboratory. The robot move path with our method and SF method in the same scenario is shown by black and yellow colors separately. We can find that when the high intensity region is in the background, SF method will extract the background as the obstacle region (the areas surrounded by red lines in the middle of Figure 9). The robot move path will be affected by these wrong regions. So, from the right half of Figure 9, we can see that the avoidance path depended on our ISRE method move in an optimal path and is obviously shorter than only depending on SF method.

## 5. Conclusions

This paper proposed a new obstacle regions 3D-perception method for mobile robots in indoor environment based on visual saliency. The mobile robot can accurately extract and perceive the obstacle in experimental environment. ISRE algorithm improved extracting effect significantly based on PCNN algorithm. When a high brightness significant area exists in the background, SF algorithm [8] will extract the

background as the saliency region. Referring to this question, we use PCNN to improve SF algorithm by simulating human visual biological processes. The saliency binarization region abstraction was fulfilled by improved PCNN multiply iteration. In order to make the ignition range more exact, we use an adaptive thresholding method to improve the ignition pulse unit of original PCNN model. To contrast the performance of our proposed method, we made some qualitative and quantitative experiments based on standard image database. The extraction of the algorithm is closer to the true value. In real scene image significant region extraction test this algorithm also showed high accuracy and further demonstrated the effectiveness and efficiency. The method is conducted on Pioneer3-DX. Experimental results have demonstrated that the presented method could locate obstacles accurately and robustly and satisfy the real-time requirement.

In a subsequent study, we will take advantage of the information about the color and texture in significant areas of the original image to recover lost significant information to enable the extraction of the real scene to be closer to



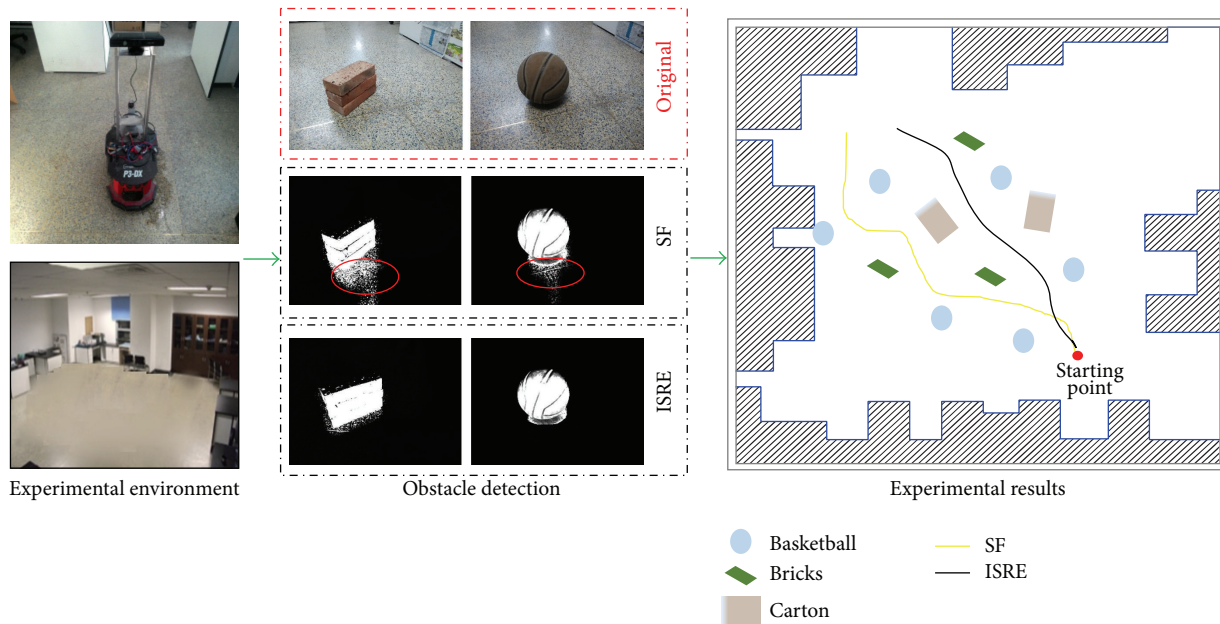


FIGURE 9: Contrast experiment results between our method and SF method.

the human visual significance. And in a further research, the algorithm's function on target recognition and tracking, obstacle detection and three-dimensional reconstruction, and the like will be considered.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgment

The research work is supported by the National Natural Science Foundation of China (61175087 and 61105033).

### References

- [1] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2376–2383, San Francisco, Calif, USA, June 2010.
- [2] A. Ohya, A. Kosaka, and A. Kak, "Vision-based navigation by a mobile robot with obstacle avoidance using single-camera vision and ultrasonic sensing," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 6, pp. 969–978, 1998.
- [3] T. Camus, D. Coombs, M. Herman et al., "Real-time single-workstation obstacle avoidance using only wide-field flow divergence," *Videre*, vol. 1, no. 3, pp. 323–330, 1999.
- [4] J. Hong, X. Tan, B. Pinette, R. Weiss, and E. M. Riseman, "Image-based homing," *IEEE Control Systems*, vol. 12, no. 1, pp. 38–45, 1992.
- [5] L. Bing, H. Bing-rong, and S. Guang, "The vision action model based on optical flow processing," *Journal of Astronautics*, vol. 24, no. 5, pp. 463–467, 2003 (Chinese).
- [6] D.-L. Dinh, M.-J. Lim, N. D. Thang, S. Lee, and T.-S. Kim, "Real-time 3D human pose recovery from a single depth image using principal direction analysis," *Applied Intelligence*, vol. 41, no. 2, pp. 473–486, 2014.
- [7] S. M. Jia, X. Zhao, W. B. Qu, and Y. Li, "Biological immune system applications on mobile robot for disabled people," *Journal of Robotics*, vol. 2014, Article ID 705198, 6 pages, 2014.
- [8] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: contrast based filtering for salient region detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 733–740, IEEE, Providence, RI, USA, June 2012.
- [9] L. Thomas and M. K. Jason, *Image Processing Using Pulse-Coupled Neural Networks*, Springer, New York, NY, USA, 2013.
- [10] G. Goswami, S. Bharadwaj, M. Vatsa, and R. Singh, "On RGB-D face recognition using Kinect," in *Proceedings of the 6th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS '13)*, pp. 1–6, IEEE, Arlington, Tex, USA, October 2013.
- [11] L. Bo, X. Ren, and D. Fox, "Depth kernel descriptors for object recognition," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 821–826, San Francisco, Calif, USA, 2011.
- [12] Y. Park, V. Lepetit, and W. Woo, "Texture-less object tracking with online training using an RGB-D camera," in *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11)*, pp. 121–126, IEEE, Basel, Switzerland, October 2011.
- [13] A. Ramey, V. Gonzalez-Pacheco, and M. A. Salichs, "Integration of a low-cost RGB-D sensor in a social robot for gesture recognition," in *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI '11)*, pp. 229–230, Lausanne, Switzerland, March 2011.
- [14] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological

- statistics,” in *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV '01)*, vol. 2, pp. 416–423, IEEE, Vancouver, Canada, July 2001.
- [15] R. Achantay, S. Hemamiz, F. Estraday, and S. Ssstrunky, “Frequency-tuned salient region detection,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1597–1604, Miami, Fla, USA, June 2009.
- [16] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952.
- [17] X. Hou and L. Zhang, “Saliency detection: a spectral residual approach,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, Minneapolis, Minn, USA, June 2007.
- [18] H. Jonathan, K. Christof, and P. Pietro, “Graph-based visual saliency,” in *Proceedings of the 20th Annual Conference on Neural Information Processing Systems*, pp. 545–552, Vancouver, Canada, 2006.
- [19] Y.-F. Ma and H.-J. Zhang, “Contrast-based image attention analysis by using fuzzy growing,” in *Proceedings of the 11th ACM International Conference on Multimedia (MULTIMEDIA '03)*, pp. 374–381, ACM, Berkeley, Calif, USA, November 2003.
- [20] R. Achanta, F. J. Estrada, P. Wils et al., “Salient region detection and segmentation,” in *Proceedings of the 6th International Conference on Computer Vision Systems*, pp. 66–75, Santorini, Greece, May 2008.
- [21] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, “Global contrast based salient region detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 409–416, IEEE, Providence, RI, USA, June 2011.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

