Saliency Region Detection Based on Region Fusion Algorithm

Haibo Lin [1, a], Qifeng Gao [1, b], Chuan Ye [2, a], Yinjun Xiong [1], cand Yanbo Wang [2, b]

[1]School of automation, CQUPT, Chongqing 400060, China;

[2]School of Advanced manufacturing engineering, CQUPT, Chongqing 400060, China.

[a]844038217@qq.com, [b]1183513828@qq.com.

Abstract

In this paper, a saliency detection algorithm based on regional fusion method is proposed to solve the problem that the current saliency detection algorithms can not accurately reflect the information of saliency. First, a simple linear iterative clustering (SLIC) method is applied to image superpixel segmentation,the superpixel region is obtained by fusion of the generated superpixel region on the premise of guaranteeing the edge information of the high quality image target. After superpixel fusion, the adjacent superpixels with similar features in the image are fused together, and the target or background in the image is gradually separated. Finally, a high quality saliency image is obtained by using the color information and the distance information of large superpixels. The experimental results show that, compared with other detection methods, this algorithm can detect the saliency target in the image more effectively, and is suitable for the industrial robot sorting system and scene recognition.

Keywords

Superpixel -segmentation;region-fusion;color contrast,saliency detection.

1. Introduction

The saliency region detection of the image is based on the human visual attention mechanism and selects the regions in the natural image that are most attractive to humans.The significance of saliency regional detection has great significance for the analysis of image content, prioritisation of important parts and rational resource allocation.Could be widely used in video detection[1],image retrieval[2],target detection[3],target recognition[4],image semantic understanding[5] and other fields.

The saliency detection algorithm can be divided into two categories as a whole.one is the method based on local contrast. The method obtains a difference by comparing an image unit with its neighboring image units in terms of gradation, color, direction, or texture, as a significant characterization.For example, Itti[6] proposed the cross-scale contrast model, which extracts features such as color, brightness and azimuth at multiple scales and then fuses to obtain the final saliency map.Ma et al.[7] used local contrast operators to calculate image saliency, but only considered the characteristics of the image's LUV color space. Zhai et al.[8] proposed a pixel-level global contrast model. However, because of the time-consuming cost, the model was only detected on the image brightness channel, and the robustness was slightly poorer.Cheng et al.[8] proposed a method based on histogram and region-based global contrast, first divide the image and then calculate the saliency value of each area,finally get a saliency map.Hou et al.[10] proposed a novel method based on spectral residues, which is statistically significant detection, it is a saliency image analysis method that belongs to a global feature.The Shannon Information Theory code thought that an image can generally be divided into sensitive and insensitive parts, that is, significant and insignificant parts of the image saliency theory.

The saliency detection method based on local contrast can detect the edge information of the image better, but it has a weaker effect on highlighting the entire saliency goal, the method of saliency detection based on global contrast is more blurred on the edge information of the target. In response to this situation, the author proposes a saliency detection method based on superpixel fusion, This method can not only preserve the target and detect the edge information through superpixel segmentation, but also ensure the complete salience target through global contrast after superpixel fusion, and finally obtain a more effective saliency map.

2. Superpixel Segmentation and Fusion

2.1 Regional Segmentation.

Region is an image processing unit between objects and pixels, it contains more information than pixels for image understanding.And the region can be easily obtained by image segmentation. In this paper, image processing is performed in the unit of region, and the original image must be divided first.In many image segmentation algorithms, the simple linear iterative clustering method was chosen in order to balance the fastness and segmentation performance. First, the image is over-divided to obtain a super-pixel region, which can maintain a good edge characteristics. Then use the similarity of the colors to merge the regions and get the segmentation map.

SLIC is a superpixel segmentation algorithm proposed by Achanta et al. [11], it uses 5-dimensional distance calculation in LAB color space.

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2}$$
$$d_{xy} = \sqrt{(x1_k - x_i)^2 + (y_k - y_i)^2} \quad (1)$$
$$D_S = d_{lab} + \frac{m}{S} d_{xy}$$

In the formula: l_k, l_i, a_k, a_i, b_k, b_i ,respectively represent two components in the LAB color space of (x_i, y_i) , (x_k, y_k), d_{lab} is the distance between two points in the LAB color space, d_{xy} is the plane distance between two points, D_s is the sum of the distance of lab distance and xy plane. M is the tightness adjustment parameter of the superpixel, here choose m=10.S is superpixel interval. We can use the parallel algorithm gSLIC in the actual calculation, so that increase the speed by more than 10 times,it processing an image with a resolution of 1280*960 within 100ms.

Information-rich superpixels can be obtained using SLIC superpixel segmentation algorithm, which for the subsequent processing.Segmentation effect as the Fig.1 shows:

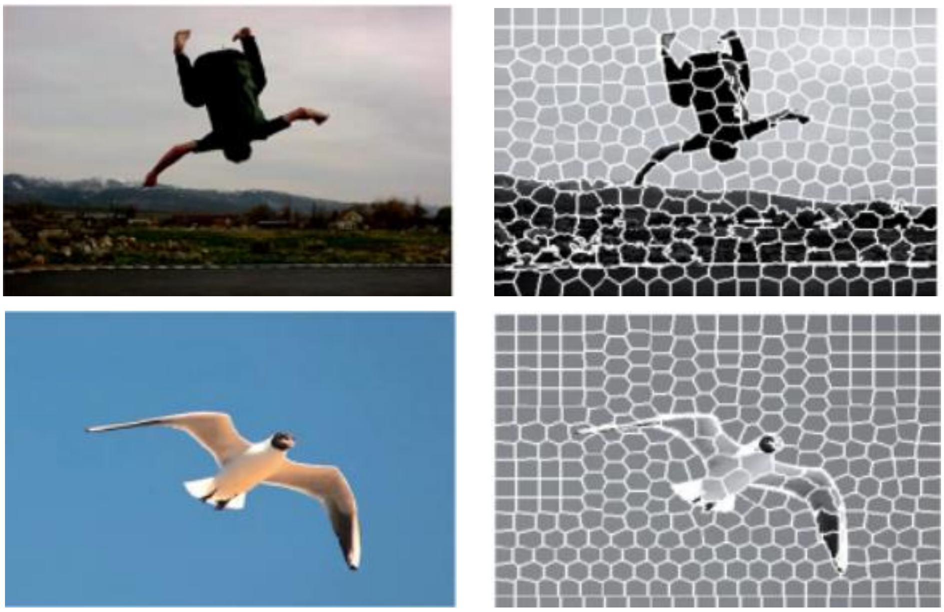

Fig 1. Superpixel segmentation

2.2 Regional Integration.

Although the superpixel can well describe the edge information of the image content, the number of superpixels is too much,if you do not deal with it, not only will it affect the speed of calculation, but it will greatly affect the accuracy and you will not be able to get the desired effect. Therefore, fusion of superpixels is the key to this algorithm.After the fusion of superpixels, not only the number of superpixels reduced to less than 20 (depending on the input image), making the follow-up processing greatly simplified, and the salient effect is more prominent, you can achieve segmentation effect nearly.

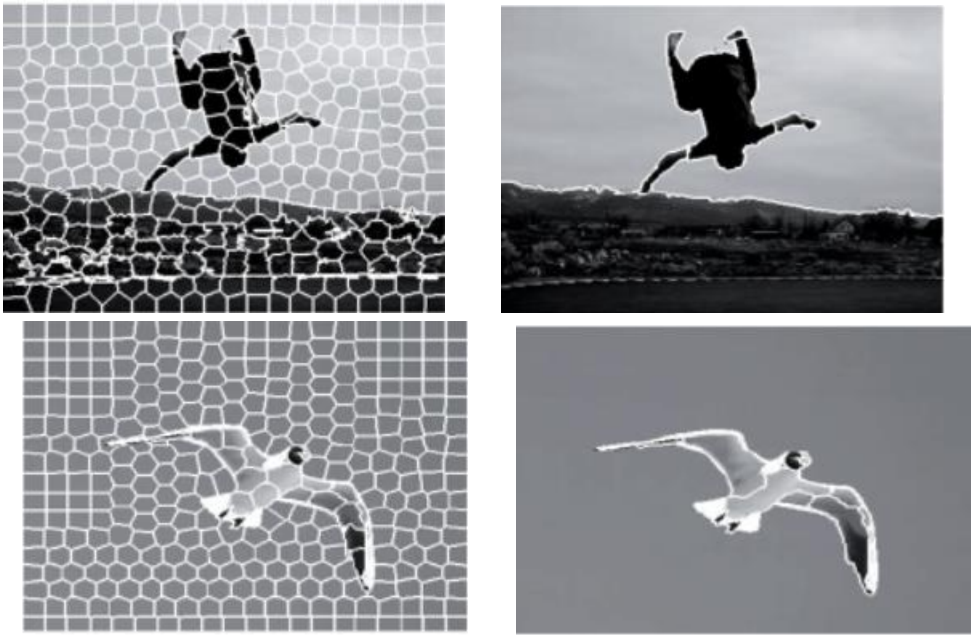

Fig 2.Superpixel fusion map

3. Saliency Region Detection

When people observe an image, they will only focus on the salience of the image according to the selective attention mechanism. The detection of salient regions is based on this selective attention mechanism. After superpixel fusion, neighboring superpixels with similar features in the image are fusioned together, and the goal or background in the image is gradually separated. At the same time, the number of superpixels has been reduced from a few hundred to a few or a dozen large superpixels containing major information,this greatly simplifies the next significant detection calculation process.

During the detection of saliency region, color contrast plays an important role. Color characteristics are the characteristics that can distinguish the salient region from the surrounding environment. This algorithm considers the color distance of a superpixel as the most important salient feature.

Define the color contrast characteristics as $\mathrm{D}_{\mathrm{color}}(r_k, r_i)$,represents the Euclidean distance between large superpixels and the LAB color space, and normalize it to [0,1].Calculate the sum of the color distances of each large superpixel and other large superpixels

$$D_c(r_k) = sum\left(\sum_{i\neq k} D_{color}(r_k, r_i)\right). \tag{2}$$

The larger of $\mathrm{D}_{\mathrm{c}}(r_k)$, the higher the significance score of the large superpixel.

In addition to color distance features, spatial distance factors between superpixels also play an important role. the space distance is also included in the significance feature for calculation. Define the space distance as $\mathrm{D}_{\mathrm{pos}}(r_k, r_i)$,represents the spatial Euclidean distance between the large superpixel and the centroid, and normalized to [0,1].Calculate the sum of the space distances between each large superpixel and other large superpixels

$$D_s(r_k) = sum(\sum_{i \neq k} D_{pos}(r_k, r_i)). \tag{3}$$

According to the above characteristic description, a comprehensive difference evaluation function is finally defined

$$D（r_k）= \frac{D_c(r_k)}{1 + cD_s(r_k)} \tag{4}$$

The parameter c in the formula is the weight of the control space distance. Therefore, according to the comprehensive difference evaluation function between the large superpixels, the formula for calculating the significance score can be given:

$$S（r_k）= 1 - \exp(-D(r_k)). \tag{5}$$

4. Experimental Results and Analysis

In order to verify the effectiveness of the proposed method, the results were compared with several commonly used visual saliency models.

The picture shows the saliency map calculated using different methods for images with inconspicuous color contrast. the saliency map obtained by this method has a clear edge of the object and the object is represented by the region,Pixels in the same region have the same salient value, the foreground part gets higher saliency, and the background section gets better suppression.

Fig 3.Saliency regions

As shown in the Fig.3, the image is more complex in the scene, especially the processing result with more texture features in the background.If you calculate the color contrast separately, it will cause a large error.the brightly colored part of the background will have a higher saliency, and this does not conform to the habit of human vision. Since this paper uses a pre-segmentation algorithm, the texture part is divided into one region,the saliency is calculated by region.This avoids the interference caused by the texture and determines the saliency of the region from a complex scene.

From the above results, it can be seen that the saliency map obtained by this method is consistent with the resolution of the original image, and can well maintain the edge features of the image region. It can maintain good performance in both simple and complex scene.

5. Conclusion

Using a superpixel fusion method for saliency region detection. By superpixel segmentation of the input image, rich superpixel information can be obtained, and then the superpixel fusion can be carried out. The number of superpixels to be processed can be greatly reduced and the image of superpixel segmentation can be obtained as a saliency target. As a basis for saliency target detection, it facilitates later processing. Using the color distance and spatial distance characteristics of the fused superpixel, the significance value is calculated, and the saliency map of the image is finally obtained. Experiments show that this method has better results compared with the more popular saliency algorithms.

References

[1] Feng Xin, Yang Dan, Zhang Ling. Saliency variation-based quality assessment for packet-loss-impaired videos. Acta Automatica Sinica,2011,37(11):1322-1331.

[2] Wang Xiang-Yang, Yang Hong-Ying, Zheng Hong-Liang, Wu Jun-Feng, A color block-histogram image retrieval based on visual weight. Acta Automatica Sinica, 2010,36 (10): 1489-1492.

[3] Ding Zheng-Hu, Yu Ying, Wang Bin, Zhang Li-Ming. Visual attention-based ship detection in multispectral imagery. Journal of Computer-Aided Design & Computer Graphics, 2011,23(3): 419-425.

[4] Wang Yan-Qing, Ma Lei, Tian Yuan. State-of-the-art of ship detection and recognition in optical remotely sensed imagery.Acta Automatica Sinica,2011,37(9):1029-1039

[5] Zhang Su-Lan,Guo Ping,Zhang Ji-Fu,Hu Li-Hua.Automatic semantic image annotation with granular analysis method. Acta Automatica Sinica,2012,38(10):688-697

[6] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1998,20 (11):1254-1259.

[7] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]//Proceedings of the 11th ACM international conference on multimedia. [S.1.]: ACM, 2003:374-381

[8] Zhai Y, Ahah M. Visual attention detection in video sequences using spatiotemporal cues[C]//Proceedings of the 14th Annual ACM International Conference on Multimedia. [S.1.]: ACM, 2006:815-824.

[9] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.1.]: IEEE,2011:409-416.

[10] Hou X. ZHANG L. Saliency detection: a spectral residual approach [C]//IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis: IEEE,2007:1-8.

[11] ACHANTA R,SHAJI A,SMITH K,et al.SLIC Superpixels [R].Lausanne:EPFL,2010.

[12] Perazzi F, Krahenbuhl P, Pritch Y, et al. Saliency filters: Contrast based filtering for salient region [C]//IEEE Computer Vision and Pattern Recognition,2012:733-740.

[13] Goferman S, Manor L Z, Tal A. Context-Aware Saliency Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(10):1915-1926.

[14] Riche N, Mancas M, Duvinage M, et al. A study of parameters affecting visual saliency assessment [J].Signal Processing: Image Communication,2013(28):642-658.

[15] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM,1981,24(6):381-395.

[16] Szeliski R. Image alignment and stitching: a tutorial [J]. Foundations and Trends in Computer Graphics and Vision,2006,2(1):1-104.

[17] Gao J, Kim S J, Brown M S. Constructing image panoramas using dual-homography warping [C]//Proceeding of IEEE Conference on Computer Vision and Pattern Recognition. New Jersey, USA,2011:49-56.

[18] Zhang F, Liu F. Parallax-tolerant image stitching [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Ohio, USA,2014:3262-3269.

[19] Lowe D G. Distinctive image features from scale-invariant key point [J]. International Journal of Computer Vision,2004,60(2):91-110.

[20] Schaefer S, McPhail T, Warren J. Image deformation using moving least squares [J]. ACM Transactions on Graphics,2006,25(3):533-540.