Image retrieval based on saliency detection in the application of the guide system

Xiaofei Sun,Wenwen Pan,Xia Wang,Wei Yuan

College of Information Science and Engineering, Zaozhuang University, Zaozhuang Shandong, 277160, China

Keywords:image segmentation; color histogram; saliency detection; saliency map; image retrieval

Abstract.A new saliency detection algorithm based on image segmentation and color histogram is proposed. The saliency values of partitions are decided by the spatial location and color difference which is measured by the difference of same color pixel number between the regions. In order to enhance salient areas and inhibit background noise a selective enhancement suppression method is used. Through the comparison of using the enhanced suppression method and not using by experiments, a better saliency map can be obtained by using this method. An image retrieval algorithm based on image saliency detection is proposed, which extracts salient areas from saliency map and divides them into blocks. Then the features of the images to search are composed of average saliency value and spatial location of each block. Finally the method is applied to guide system, which can quickly and accurately retrieve tourism information.

Introduction

More and more people like tourism nowadays, tourists carrying hand-held photography equipment can conveniently capture images. Obtain the corresponding spots information using these images will be a convenient guide mode, and the use of image retrieval technology can achieve this purpose. Image search method based on content is the focus of current research, the latest research tend to integrate multi features. One is to compute the similarity of each feature and then compute the weighted sum[1,2,3], so allocation of weights is of strong subjectivity, another is to combine several features to calculate similarity, namely constructing hybrid feature[4,5,6,7,8]. Literature[9] proposed a method to structure hybrid features using shape and color based on multiple integral, and derived a group of features invariant to affine transformation of shape and diagonal shift transformation of color. Literature[10] proposed a new image segmentation method which segmented image into regions having same visual consistency and extracted features of these regions for image retrieval, and this method obtained good effect when experiment on images including objects of similar color and similar size, but the result is poor when experiment on images including objects of different colors and similar shape. A new image retrieval method based on image saliency detection is proposed in this paper and applied to the guide system.

Saliency detection as an important means to acquire the main contents of image has been widely used in many aspects, such as image cropping[11], image and video compression[12], image retrieval [2,3,4,5,6], medical diagnosis, picking robot etc. Although the latest research results of saliency detection method have been obtained some deficiencies still exists. Literature [13] proposed a saliency detection method which divided images into 8 * 8 blocks firstly, next computed Local feature and Regional feature of each block in multiple scales, and then combined with the salient object edge to obtain the saliency map. To some extent, this method improves hollow effect that the local feature brought in, but distinction between salient object and the surrounding background is not obvious. Literature [14] constructed sparse color histogram based on image segmentation, and then calculated saliency value of each region. This method does not bring in the hollow effect but need to construct sparse histogram, and in the calculation of color distance between regions is very complex, so it costs longer processing time.

To meet the requirements of efficiency and effectiveness, a new saliency detection method is proposed in this paper. First, use the difference between the pixel numbers of same color in

different regions as the color dissimilarity of them, coordinated with the spatial feature to extract saliency value. Then enhance salient regions and inhibit background noise selectively to solve the problem that all pixels in the same region with same saliency value. Finally apply the saliency detection method to guide system to retrieve images.

Saliency detection based on image segmentation

This paper uses graph-based approach for image segmentation [15], and in order to reduce the cost of calculation quantify each component of each region for 12 values in RGB space, so the total number of colors is reduced to 1728[7]. In order to describe the saliency degree of the whole image use the difference between the pixel numbers of same color in different regions as the color dissimilarity of them. Then establish the color histogram of each region, the color distance between r_m and r_n is expressed as follows,

$$D_c(r_m, r_n) = \lambda \sum_{i=1}^{N_c} |N_{m,i} - N_{n,i}| i \tag{1}$$

Among them i is the sequence number of colors being from the calculation of the RGB component，$i = k^2 B + kG + R, (k = 12)$。 $N_c =$ 1728 is the total number of colors. $N_{m,i}$, $N_{n,i}$ are respective the number of the i-th color in region r_m, r_n. λ is the adjustment parameter when taken the value of 0.0001 it works best. The number difference of pixel with i-th color in regions r_m, r_n is multiplied by the sequence number i, which is the dissimilarity of i-th color between two regions. The accumulation of dissimilarities of all colors in regions r_m, r_n is the total color dissimilarity, $D_c(r_m, r_n)$, between them. The reason why we do this is that the number difference of pixel with different color may be same, multiplied by the sequence number of this color can distinguish each color's contribution to the total color dissimilarity between regions. And because the bigger the spatial distance is, the smaller the effect on each other is. Therefore, the saliency degree of each region in the image is defined as follows,

$$S(r_i) = \sum_{r_k \neq r_i} \frac{1}{1 + \delta D_s(r_k, r_i)} w(r_k) D_c(r_k, r_i) \tag{2}$$

Using $\dfrac{1}{1 + \delta D_s(r_k, r_i)}$ as the weight of distance, where $D_s(r_k, r_i)$ is the city block distance between the centroids of regions r_k, r, city block distance used here is more efficient than the Euclidean distance, and δ is a adjusting parameters and set to 200. $w(r_k)$ is the number of pixels in region r_k and used to adjust the influence of other regions with different size to region r_i,

$$S(r_k)' = S(r_k) \frac{1}{1 + \delta_2 D_o(r_k)} \tag{3}$$

Considering the tourists often make the target placed in the centre of image when take photos, we added the weight $\dfrac{1}{1 + \delta_2 D_o(r_i)}$ to the saliency value of each region. The bigger the distance $D_o(r_i)$ between r_i and center of the image is, the smaller the saliency value is. δ_2 is set to10 to enhance the effect of spatial distance.

Experiment on the MSRA Salient Object Database which contains all kinds of 5000 images with this method can both obtain good saliency map whether the foreground and background contrast is obvious or less obvious. The results are shown in Fig. 1.
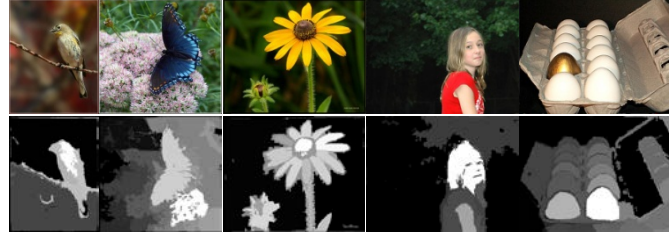
Fig. 1 the original image and the saliency map extracted by this method

Selective enhancement and inhibition

Because this algorithm based on image segmentation, the same region was given the same saliency value, lacking in detail display. Selective enhancement and inhibition in this paper makes the improvement that the internal details of region become more detailed while the background becomes suppressed. The saliency value of point (x, y) defined as follows,

$$S_{(x,y)} = \rho S(r_k)' G_{(x,y)} \tag{4}$$

$$G_{(x,y)} = \max(r_{(x,y)}, g_{(x,y)}, b_{(x,y)}) \tag{5}$$

$S(r_k)'$ is the saliency value of region r_k, $G_{(x,y)}$ is the maximum of RGB component at the point of (x, y).

Comparisons of saliency maps

Experiments were conducted on MSRA database and saliency maps extracted by several typical methods and this method were compared as shown in Fig. 2.

The images in Fig.2 from left to right are respective the original image, the saliency maps extracted by Itti[16], Ma and Zhang[17], Harel et al. [18], Hou and Zhang [19], Achanta et al. [20], this method and the improved method with selective enhancement and inhibition. Only a small amount of salient pixels are obtained by the IT method, MZ and SR methods only have a good description of the profile, GB could bring fuzzy effect, AC could not distinguish salient region and background obviously. While this method can better highlight the salient region, especially the improved method can generate saliency map with internal details more detailed and the background noise further suppressed.
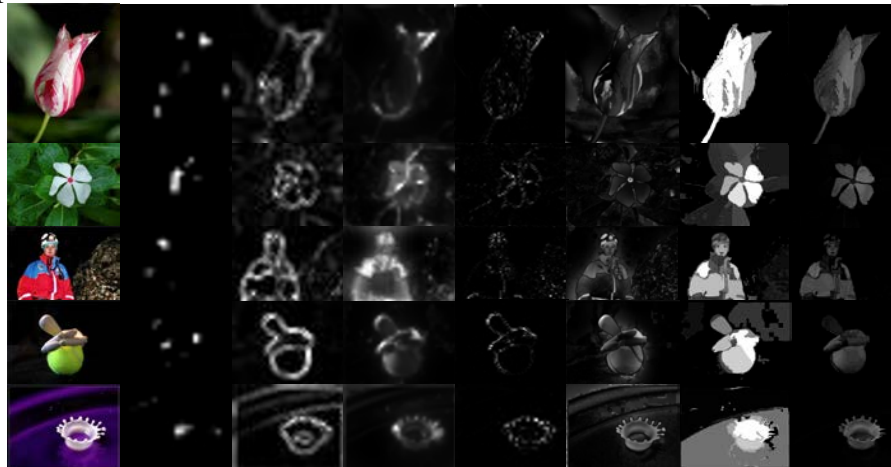
(a)original (b)IT (c)MZ (d)GB (e)SR (f)AC (g)our (h) our
Fig. 2 Comparisons of saliency maps

Image retrieval based on salient region

Image search based on salient region is a kind of content-based image retrieval, can eliminate the background interference. The saliency maps used for retrieval are generated with the method

proposed in this paper, and steps for image retrieval are as follows,

(1) Extract the salient region from saliency map, and divide it into 10×10 blocks. The selection of the number blocks is in order to achieve the balance between the efficiency and the accuracy.

(2) The average saliency and spatial location of each block are stored as the feature information used for retrieval.

(3) Calculate the feature information of image to-retrieve when go to search images and compare with the feature information of each image in the image library. The comparison strategy is that calculate the saliency differences between corresponding blocks in both images combined with spatial weights to obtain the whole dissimilarity between the two images.

(4) Results are sorted according to the whole dissimilarity of images, and the less it is the more similar to to-retrieve image.

Read the saliency value of pixels from the upper boundary of the saliency map line by line, if the saliency value of all pixels in this line less than a threshold T_s, then continue to traverse the next line until saliency value of pixel greater than the threshold is found and remark this line as the upper boundary of regions. The other three boundaries of salient region can be gotten by the same way. Fig. 3 shows the extracted salient region according to this method. The salient region divided into 10×10 blocks $(u_1, u_2, u_3, \cdots, u_{100})$ can retain salient object in proportion.
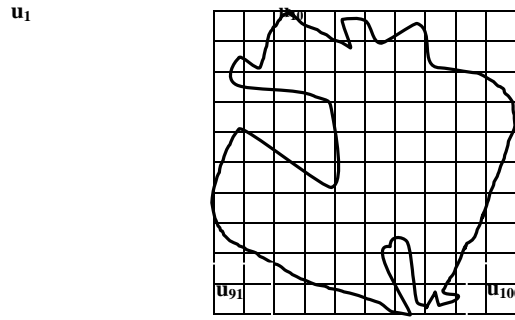

Fig. 3 Schematic diagram of salient region segmentation

The average saliency of block u_i is the result that the saliency value of each pixel is added up and divided by the number of pixels in block u_i. The average saliency and sequence number i of block u_i is saved as the feature for image retrieval.

When retrieve images, process images to-retrieve with the method above to obtain the feature information used for retrieval. The dissimilarity between the two images is calculated as follows,

$$D(I_m, I_n) = \sum_{i=1}^{N_u} \frac{i}{N_u} D_s(u_{(m,i)}, u_{(n,i)}) \tag{6}$$

$D_s(u_{(m,i)}, u_{(n,i)})$ is average saliency difference of u_i located in I_m, I_n respectively. Considering the spatial relationship, we add $\frac{i}{N_u}$ as a linear distance weight，where $N_u = 100$ is the number of blocks, i (1,2,3…100)is the sequence number of block.Sort the result according to the $D(I_m, I_n)$, the less the value is the more similar the image is.

Applied in guide system

Development of this project is dedicated to the ancient city of Taierzhuang tour guide system, the guide function of the system are designed as follows: automatic receiving and playing self-help tour guide, automatic recognition of scenic spots, automatic directing way by voice, tourists real-time location tracking, tourist statistics, the crowd guidance. The recognition of scenic spots is the premise of self-help guide function, so this paper takes it as the research object.

The tourists usually place the scenery in the center of image when take photos, which is in favor of obtaining the interested object through the saliency map. Saliency detection is applied to guide system in this paper, and tourists use the photo they taken to retrieve in the image feature database, so as to confirm the names of scenic spots and obtain information feedback to realize the functions of tour guide. In this way tourists can obtain self-help guide function at the time of taking photos, which brings great convenience for tourists.

This paper chose scenic spots of Taierzhuang city for matching experiment, select scenic spots in the ancient city to build image database. Taking into account the tourists taking photos differ from man to man, so the take photos at different angles, different distance in order to improve the matching accuracy. Then extract the feature information of scenic spots to establish a retrieval feature library. This image retrieval method is applied to the tour guide system for scenic spots recognition, and the corresponding images can be ranked first, which prove that this method used for self-help guides and scenic spot introduction has very good prospect.

Conclusion

A new saliency detection method is proposed in this paper first. Use a graph-based method for image segmentation, establish color histogram for each region and use pixel number difference of same color as measurement to calculate the saliency of region. Because graph-based method for image segmentation will lead in a result that one region has one saliency, a selective enhancement and suppression method is adopted to make the saliency map better. Compared with other saliency detection method, the effect of the improved method is very good. Then an image retrieval method based on salient map is proposed. Extract the salient region, divide the salient region into 10×10 blocks, and use average saliency and position of each block as retrieval feature. Through experiments prove that this kind of image retrieval method has a good effect when applied to guide system, the effect is very good. Since the rotation angle in image guide system is very small in general, the image retrieval without considering the rotation factors. The next step is to study the rotational invariants and mixed invariants in image retrieval, and apply the saliency detection method in image compression, pedestrian detection and other aspects.

Acknowledgments

References

[1] Piamsa-nga P, Alexandridis N A, Srakaew S, et al. Multi-feature content based image retrieval[C]//Proceedings of International Conference on Computer Graphics and Imaging.Halifax,Canada:IASTED,1998.
[2] Sun J D, Cui J T, Wu X S, et al. Color image retrieval based on color and shape features[J]. Journal of Image and Graphics, 2004,9(7):820-827.
[3] Deng C Q, Feng G. Content-based image retrieval using combination features[J]. Computer Applications, 2003,23(7):100-102.
[4] Wu J,Qiu Z D. An overview: low-level feature fusion in content-based image retrieval[J]. Journal of Image and Graphics,2008,13(2):189-197.
[5] Hu M. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory, 1962,(8):179-187.
[6] Wang Y P, Huang X S, Li X L, et al. Moment invariants of gray distorted based on Hu moments[J]. Journal of Dalian Maritime University, 2009, 34(4):23-27.

[7] Mingdru F, Tuytelaars T, Gool L V, et al. Moment invariants for recognition under changing viewpoint and illumination[J].Computer Vision and Image Understanding,2004,94(1):3-27.

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

[9] Gong Ming, Gao Weiguo, Li Hua. Application of combined shape and color invariants to image retrieval.Journal of image and Graphics[J],2013,18(8):
990-1003.

[10] He Guangnan, Yang Yubin, Ruan Jiabin, Lin Jinjie.Image retrieval based on visual consistency. Journal of image and Graphics.2011,16(4):503-509.

[11] Hua Shun-Gang, Chen Guo-Peng, Shi Shu-Sheng. Image resizing algorithm based on similarity criterion. Computer Engineering, 2012, 38(4): 191-193.

[12] Gupta R, Chaudhury S. A scheme for attentional video compression[J]. Pattern Recognition and Machine Intelligence,2011, 6744: 458-465.

[13] GUO Ying-Chun, YUAN Hao-Jie,WU Peng.Image Saliency Detection Based on Local and Regional Features[J]. ACTA AUTOMATICA SINICA.2013,39(8):1214-1224.

[14] Ming-Ming Cheng, Guo-Xin Zhang,et al.Global Contrast based Salient Region Detection[C]. IEEE International Conference on Computer Vision and Pat-tern Recognition(IEEE CVPR)), 2011.409-416.

[15]P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation[J]. International Journal of Computer Vision, 2004,59(2):167–181.

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

[17] Y.-F. Ma and H.-J. Zhang. Contrast-based image attention analysis by using fuzzy growing[C]. In: Proceedings of the 11th ACM International Conference on Multimedia. New York:ACM, 2003. 374-381.

[18] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems [C]. 2007,19:545-552.

[19] X. Hou and L. Zhang. Saliency detection: A spectral residual approach[C]. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis,MN: IEEE, 2007. 1-8.

[20] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk. Salient region detection and segmentation[C]. International Conference on Computer Vision Systems 2008.66-75.