Salient Regions Detection from Images

Kadam Vijay Murlidhar
Computer Science and Engineering, Siddhi Vinayak College of Science & Higher Education,
Alwar, India.
vijaykadama@gmail.com

Atul Gaur
Computer Science and Engineering, Siddhi Vinayak College of Science & Higher Education,
Alwar, India.

Abstract—Discovery of outwardly remarkable picture districts is helpful for applications like protest division, versatile pressure, and question acknowledgment. In this paper, we present a technique for remarkable locale recognition that yields full determination saliency maps with very much characterized limits of notable items. These limits are safeguarded by holding generously more recurrence content from the first picture than other existing methods. Our strategy abuses highlights of shading and luminance, is easy to actualize, and is computationally productive. We contrast our calculation with five best in class striking district location techniques with a recurrence area investigation, ground truth, and a remarkable question division application. Our strategy beats the five calculations both on the ground-truth assessment and on the division errand by accomplishing both higher exactness and better review.

Keywords—Image Segmentation, Salient Image Regions, BOV-Bag Of Visual Word.

I. INTRODUCTION

Visual saliency is the perceptual quality that makes a question, individual, or pixel emerge in respect to its neighbors and hence catch our consideration. Visual consideration comes about both from quick, pre-mindful, base up visual saliency of the retinal contribution, and also from slower, top-down memory and volition based handling that is assignment subordinate [24]. The concentration of this paper is the programmed location of outwardly notable locales in pictures, which is valuable in applications, for example, versatile substance conveyance [22], versatile district of-intrigue based picture pressure [4], picture division [18, 9], question acknowledgment [26], and content mindful picture resizing [2]. Our calculation discovers low-level, pre-mindful, base up saliency. It is roused by the natural idea of focus encompass differentiate, however does not depend on any organic model.

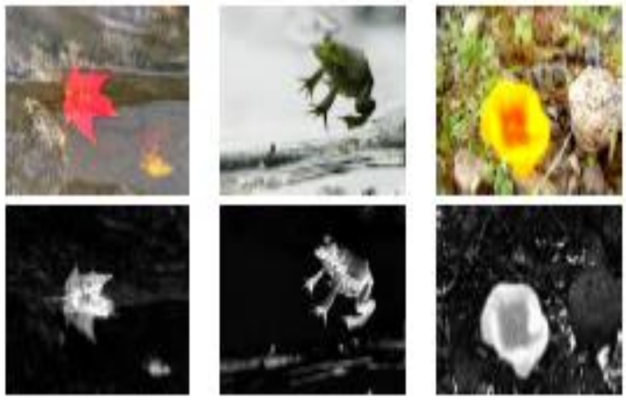

Fig.1 .Original Images and Their Saliency Maps Using Our Algorithm

Current techniques for saliency identification produce locales that have low determination, ineffectively characterized outskirts, or are costly to register. Furthermore, a few techniques create higher saliency esteems at question edges as opposed to producing maps that consistently cover the entire protest, which comes about from neglecting to abuse all the spatial recurrence substance of the first picture. We examine the spatial frequencies in the first picture that are held by five conditions of-the art systems, and outwardly show that these procedures basically work utilizing to great degree low-recurrence content in the picture. We acquaint a recurrence tuned approach with assess focus encompass differentiate utilizing shading and luminance includes that offers three focal points over existing techniques: consistently featured remarkable locales with well-defined limits, full determination, and computational productivity. The saliency outline can be all the more successfully utilized as a part of numerous applications, and here we exhibit comes about for protest division. We give a target correlation of the exactness of the saliency maps against five state-of the-craftsmanship techniques utilizing a ground truth of a 1000 pictures. Our strategy beats these strategies regarding exactness furthermore, review.

II. GENRAL APPROCHES TO DETERMINE SALIENCY

The term saliency was utilized by Tsotsos et al. [27] and Olshausen et al. [25] in their work on visual consideration, and by Itti et al. [16] in their work on quick scene investigation. Saliency has likewise been alluded to as visual consideration [27, 22], capriciousness, irregularity, or shock [17, 14]. Saliency estimation strategies can extensively be named organically based, simply computational, or a blend. When all is said in done, all strategies utilize a low-level approach by deciding complexity of picture districts with respect to their environment, utilizing at least one highlight of force, shading, and introduction. Itti et al. [16] construct their technique with respect to the naturally conceivable engineering proposed by Koch and Ullman [19]. They decide focus encompass differentiate utilizing a Difference of Gaussians (DoG) approach. Frintrop et al. [7] exhibit a technique roused by Itti's strategy, however they figure center surround contrasts with square channels and utilize essential pictures to accelerate the computations. Different strategies are absolutely computational [22, 13, 12, 1] and are not founded on natural vision standards. Mama and Zhang [22] and Achanta et al. [1] appraise saliency utilizing focus encompass include separations. Hu et al. [13] gauge saliency by applying heuristic measures on beginning saliency measures acquired by histogram thresholding of highlight maps. Gao and Vasconcelos [8] boost the common data between the element dispersions of focus and encompass areas in a picture, while Hou and Zhang [12] depend on recurrence space preparing. The third classification of strategies are those that fuse thoughts that are mostly in view of natural models and halfway on computational ones. For example, Harel et al. [10] make include maps utilizing Itti's strategy yet play out their standardization utilizing a diagram based approach. Different techniques utilize a computational approach like augmentation of data [3] that speaks to a naturally conceivable model of saliency identification. A few calculations distinguish saliency over various scales [16, 1], while others work on a solitary scale [22, 13]. Additionally, singular element maps are made independently and after that joined to acquire the last saliency delineate, 22, 13, 7], or a component consolidated saliency outline specifically got [22, 1].

III. SPATIAL FREQUENCY CONTENT OF SALIENCY MAPS

To investigate the properties of the five saliency calculations, we inspect the spatial recurrence content from the first picture that is held in figuring the last saliency delineate. It will be appeared in Sec. 4.3 that the scope of spatial frequencies held by our proposed calculation is more suitable than the calculations utilized for examination. For effortlessness, the accompanying examination is given in one measurement and expansions to two measurements are elucidated when essential. In technique IT, a Gaussian pyramid of 9 levels (level 0 is the first picture) is worked with progressive Gaussian obscuring and down sampling by 2 in each measurement. On account of the luminance picture, these outcomes in a progressive diminishment of the spatial frequencies held from the information picture. Each smoothing operation roughly parts the standardized recurrence range of the picture. Toward the finish of 8 such smoothing operations, the frequencies held from the range of the first picture at level 8 territory inside [0, Π/256]. The procedure processes contrasts of Gaussian-smoothed pictures from this pyramid, resizing them to size of level 4, which brings about utilizing recurrence content from the first picture in the range [Π/256, Π/16]. In this recurrence go the DC (mean) segment is evacuated alongside roughly 99% ((1- 1/ $16^{2)}$*100) of the high frequencies for a 2-D picture. All things considered, the net data held from the first picture contains not very many subtle elements furthermore, speaks to an exceptionally foggy variant of the first picture (see the band-pass sifted picture of Fig. 2(b)). In technique MZ, a low-determination picture is made by averaging pieces of pixels and after that downsampling the separated picture with the end goal that each square is spoken to by a solitary pixel having its normal esteem. The averaging operation performs low-pass sifting. While the creators don't give a piece size to this operation, we got great with a square size of 10_10 pixels, and all things considered the frequencies held from the first picture are in the range [0, Π/10]. In strategy GB, the underlying strides for making highlight maps are like IT, with the distinction that less levels of the pyramid are utilized to discover focus encompass contrasts. The spatial frequencies held are inside the range [Π/128, Π/8]. Around 98% ((1- 1/ 8^2) * 100) of the high frequencies are disposed of for a 2D picture. As delineated in Fig. 2(d), there is somewhat higher recurrence content than in 2(b). In strategy SR, the info picture is resized to 64 * 64 pixels (through low-pass sifting and downsampling) in light of the contention that the spatial determination of pre-mindful vision is exceptionally constrained. The subsequent recurrence substance of the resized picture in this way shifts as per the first size of the picture. For instance, with input pictures of size 320 * 320 pixels (which is the inexact normal measurement of the pictures of our test database), the held frequencies are restricted to the range [0, Π/5]. As found in Fig. 2(e), higher frequencies are smoothed out. In technique AC, a distinction of-implies channel is utilized to appraise focus encompass differentiate. The most minimal frequencies held rely upon the measure of the biggest encompass channel (which is half of the picture's littler measurement) and the most noteworthy frequencies rely upon the span of the littlest focus channel (which is one pixel). All things considered, technique AC viably holds the whole scope of frequencies (0, Π] with an indent at DC. All the high frequencies from the first picture are held in the saliency delineate not every single low recurrence (see Fig. 2(f)).

IV. SEGMENTATION BY FIXED THRESHOLDING

For a given saliency delineate, saliency esteems in the range [0; 255], the most straightforward approach to acquire a parallel veil for the notable question is to limit the saliency outline an edge Tf inside [0; 255]. To think about the nature of the diverse saliency maps, we differ this limit from 0 to 255, and process the exactness and review at each estimation of the edge. The subsequent accuracy versus review bend is appeared in Fig. 2. This bend gives a solid examination of how well different saliency maps feature notable areas in pictures. It is intriguing to take note of that Itti's strategy indicates high exactness for a low review (< 0:1), and after that the precision drops steeply. This is on the grounds that the notable pixels from this strategy fall well inside striking areas and has close uniform esteems; however don't cover the whole remarkable protest. Techniques GB and AC have comparative execution in spite of the way that the last creates full determination maps as yield. At most extreme review, all techniques have similar low accuracy esteem. This occurs at edge zero, where all pixels from the saliency maps of every strategy are held as positives, prompting an equivalent incentive for genuine and false positives for all techniques.

V. SEGMENTATION BY ADAPTIVE THRESHOLDING

Maps created by saliency indicators can be utilized in remarkable protest division utilizing more refined strategies than straightforward thresholding. Saliency maps created by Itti's approach have been utilized as a part of unsupervised protest division. Han et al. [9] utilize a Markov arbitrary field to incorporate the seed esteems from Itti's saliency outline with low-level highlights of shading, surface, and edges to develop the remarkable question areas. Ko and Nam [18] use a Support Vector Machine prepared on picture fragment highlights to choose the remarkable districts of enthusiasm utilizing Itti's maps, which are then grouped to separate the notable items. Mama and Zhang [22] utilize fluffy developing on their saliency maps to limit remarkable areas inside a rectangular locale. We utilize a less complex technique for portioning remarkable articles, which is an altered adaptation of that displayed in [1]. Their strategy makes utilization of the force and shading properties of the pixels alongside their saliency esteems to section the protest. Considering the full determination saliency outline, method over-fragments the information picture utilizing k-implies bunching and holds just those portions whose normal saliency is more noteworthy than a consistent limit. The parallel maps speaking to the notable protest are along these lines acquired by allocating ones to pixels of picked fragments and zeroes to whatever is left of the pixels. We make two changes to this technique. To begin with, we supplant the slope climbing based k-implies division calculation by the mean-move division calculation [5], which gives better division limits. We perform mean-move division in Lab shading space. We utilize settled parameters of 7, 10, 20 for sigmaS, sigmaR, and minRegion, separately, for every one of the pictures (see [5]). We additionally present a versatile limit that is picture saliency subordinate, rather than utilizing a steady edge for each picture. This is like the versatile limit proposed by Hou and Zhang [12] to recognize proto-objects.

The versatile edge (Ta) esteem is resolved as two times the mean saliency of a given picture:

$$Ta = \frac{2}{(W*H)} \sum_{a=1}^{W-1} \sum_{y=0}^{H-1} s\,(x,y) \qquad \ldots\ldots 1$$

Where W and H are the width and stature of the saliency outline pixels, individually, and S(x; y) is the saliency estimation of the pixel at position (x; y). A couple of consequences of remarkable question division utilizing our changes are appeared in Fig. 2. Utilizing this changed approach, we acquire binarized maps of notable question from each of the saliency calculations. Normal estimations of exactness, review, and F-Measure (Eq. 10) are acquired over a similar ground-truth database utilized as a part of the past trial.

$$F_\beta = \frac{(1+\beta^2\ )\ Precision * Recall}{\beta^2 * Precision + Recall} \qquad \ldots\ldots\ldots\ldots 2$$

We utilize _2 = 0:3 in our work to measure exactness more than review. The examination is appeared in Fig. 2. Itti's strategy (IT) demonstrates a high exactness however extremely poor showed, that it is more qualified for look following analyses, yet maybe not appropriate for remarkable question division. Among all the techniques, our strategy (IG) demonstrates the most noteworthy accuracy, review, and F_ esteems. Our technique plainly beats substitute, condition of-the art calculations. Nonetheless, similar to all saliency identification strategies, it can come up short if the question of intrigue is not unmistakable from the foundation as far as visual difference (see Fig 2(b), first line.

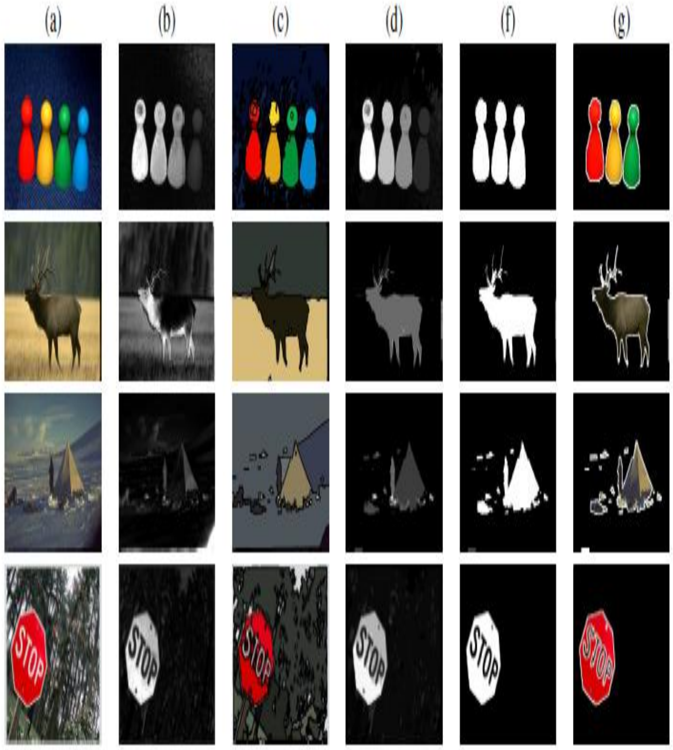

Figure 2: Visual comparison of saliency maps. (a) original image, (b) saliency maps using the method presented by, Itti [16], (c) Ma andZhang [22], (d) Harel et al. [10], (e) Hou and Zhang [12], (f) Achanta et al. [1], and (g) our method. Our method generates sharper anduniformly highlighted salient regions as compared to other methods.

VI. CONCLUSION

We played out a recurrence space investigation on five state of-the-workmanship saliency techniques, and thought about the spatial recurrence content held from the first picture, which is then utilized as a part of the calculation of the saliency maps. This investigation delineated that the lacks of these strategies emerge from the utilization of a wrong scope of spatial frequencies. In light of this investigation, we introduced a recurrence tuned approach of registering saliency in pictures utilizing low level highlights of shading and luminance, which is anything but difficult to actualize, quick, and gives full determination saliency maps. The subsequent saliency maps are more qualified to notable question division, exhibiting both higher accuracy and preferred review over the five best in class strategies.

REFERENCES

[1]. R. Achanta, F. Estrada, P. Wils, and S. S¨usstrunk. Salient region detection and segmentation. International Conference on Computer Vision Systems, 2008.
[2]. S. Avidan and A. Shamir. Seam carving for content-aware image resizing. ACM Transactions on Graphics, 26(3), 2007.
[3]. N. Bruce and J. Tsotsos. Attention based on information maximization. Journal of Vision, 7(9):950–950, 2007.
[4]. C. Christopoulos, A. Skodras, A. Koike, and T. Ebrahimi. The JPEG2000 still image coding system: An overview.IEEE Transactions on Consumer Electronics, 46(4):1103– 1127, 2000.
[5]. C. Christoudias, B. Georgescu, and P. Meer. Synergism in low level vision. IEEE Conference on Pattern Recognition,2002.
[6]. J. L. Crowley, O. Riff, and J. H. Piater. Fast computation of characteristic scale using a half octave pyramid. International Conference on Scale-Space theories in Computer Vision, 2003.
[7]. S. Frintrop, M. Klodt, and E. Rome. A real-time visual attention system using integral images. International Conference on Computer Vision Systems, 2007.
[8]. D. Gao and N. Vasconcelos. Bottom-up saliency is a discriminant process. IEEE Conference on Computer Vision, 2007.
[9]. J. Han, K. Ngan, M. Li, and H. Zhang. Unsupervised extraction of visual attention objects in color images. IEEE Transactions on Circuits and Systems for Video Technology, 16(1):141–145, 2006.
[10]. J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems,19:545–552, 2007.
[11]. S. S. Hemami and T. N. Pappas. Perceptual metrics for image quality evaluation. Tutorial presented at IS&T/SPIE Human Vision and Electronic Imaging, 2007.
[12]. X. Hou and L. Zhang. Saliency detection: A spectral residual approach. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[13]. Y. Hu, X. Xie, W.-Y. Ma, L.-T. Chia, and D. Rajan. Salient region detection using weighted feature maps based on the human visual attention model. Pacific Rim Conference on Multimedia, 2004.
[14]. L. Itti and P. F. Baldi. Bayesian surprise attracts human attention. Advances in Neural Information Processing Systems, 19:547–554, 2005.
[15]. L. Itti and C. Koch. Comparison of feature combination strategies for saliency-based visual attention systems. SPIEHuman Vision and Electronic Imaging IV, 3644(1):473–482,1999.
[16]. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254– 1259, 1998.
[17]. T. Kadir, A. Zisserman, and M. Brady. An affine invariant salient region detector. European Conference on Computer Vision, 2004.
[18]. B. C. Ko and J.-Y. Nam. Object-of-interest image segmentation based on human attention and semantic region clustering. Journal of Optical Society of America A, 23(10):2462– 2470, 2006.
[19]. C. Koch and S. Ullman. Shifts in selective visual attention: Towards the underlying neural circuitry. Human Neurobiology, 4(4):219–227, 1985.
[20]. T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[21]. D. G. Lowe. Distinctive image features from scale-invariant feature points. International Journal of Computer Vision, 60:91–110, 2004.
[22]. Y.-F. Ma and H.-J. Zhang. Contrast-based image attention analysis by using fuzzy growing. In ACM International Conference on Multimedia, 2003.
[23]. D. Marr. Vision: a computational investigation into the human representation and processing of visual information. W.H. Freeman, San Francisco, 1982.
[24]. E. Niebur and C. Koch. The Attentive Brain, chapter Computational architectures for attention, pages 163–186. Cambridge MA:MIT Press, October 1995.
[25]. B. Olshausen, C. Anderson, and D. Van Essen. A neurobiological model of visual attention and invariant pattern recognition Neuroscience, 13:4700–4719, 1993.
[26]. U. Rutishauser, D. Walther, C. Koch, and P. Perona. Is bottom-up attention useful for object recognition? IEEE Conference on Computer Vision and Pattern Recognition, 2, 2004.
[27]. J. K. Tsotsos, S. M. Culhane,W. Y. K.Wai, Y. Lai, N. Davis, and F. Nuflo. Modeling visual attention via selective tuning. Artificial Intelligence, 78(1-2):507–545, 1995.
[28]. Z. Wang and B. Li. A two-stage approach to saliency detection in images. IEEE Conference on Acoustics, Speech and Signal Processing, 2008.