International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 06 | June -2017 www.irjet.net p-ISSN: 2395-0072

A REVIEW ON IMPLEMENTATION OF HIGH DIMENSION COLOUR TRANSFORM IN DOMAIN OF IMAGE PROCESSING

Anchal A. Khadse[1], Renu G. Dhokte[2], Mr. M. H. Nerkar[3]

[1,2] M.Tech student, Electronics and Telecommunication Engineering, Government College of Engineering, Jalgaon, Maharashtra, India
[3]Professor in Electronics and Telecommunication Engineering, Government College of Engineering, Jalgaon, Maharashtra, India

---

Abstract - *Extraction of salient region from colour image is very useful nowadays so this topic deals with the study of salient region detection in an image with the help of high dimension colour transform algorithm. To extract salient region from the image it is necessary to design saliency map. For computation of saliency map, global and local features of the image are needed to find out. The creation of saliency map is a first approach of this work, and it is a linear combination of colours in high dimensional colour space. By mapping the low-dimensional red, green, and blue colour to a feature vector in a high-dimensional colour space, we show that we can composite an accurate saliency map by finding the optimal linear combination of colour coefficients in the high-dimensional colour space. The performance of our saliency estimation is improved by our second approach which is used to utilize relative location and colour contrast between superpixels as features and to resolve the saliency estimation from a trimap via a learning-based algorithm. We here used number of images as a trained dataset for analysis of different parameters in colour images.*

Key Words: **Salient Region; Superpixels; trimap; feature vector; HDCT method; Learning based algorithm.**

1. INTRODUCTION

Very informative portion in an image is termed as salient region. For detection and analysis of that region there are various algorithms which were proposed until now. But HDCT algorithm detect salient region of image in terms of saliency map with the help of global and local feature extraction. Many previous studies have shown that salient region detection is useful, and it has been applied to many applications including segmentation [1], object recognition [2], image retargeting, photo collage, image quality assessment, image thumb nailing, and video compression. The concepts of human visual perception improvise the development of salient region detection in the domain of image processing. As color is a very important visual cue to human, many salient region detection techniques are built upon distinctive color detection from an image. This project contains an approach to automatically detect salient regions in an image. First approach is to estimate the approximate locations of salient regions by using a tree-based classifier. A tree-based classifier classifies each superpixel as a foreground, background or unknown pixels. This classifier classifies those foreground and background regions into salient and non-salient regions with high confidence. The unknown region is classified with low confidence but that unknown region includes ambiguous features in it. All the three regions foreground, background as well as unknown regions forms an initial trimap, and our purpose is to resolve the ambiguousness from that unknown regions to calculate accurate saliency map. From the trimap, we can conclude to use two different methods, high-dimensional color transform (HDCT)-based method and local learning based method to estimate the saliency map. The results of these two methods will be combined together to form our final saliency map. Fig. 1 shows examples of our saliency map and salient regions from trimap.

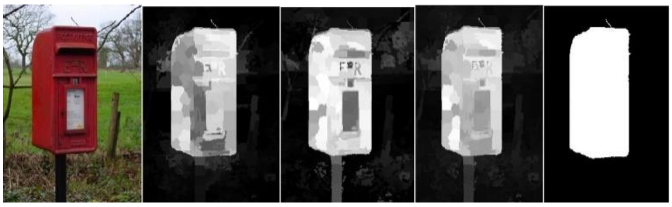


Fig -1: Example of our salient region detection method from a trimap. (a) Input. (b) Trimaps. (c) Saliency maps. (d) Salient region with final segmentation. (d) Ground truth.

2. LITERATURE REVIEW

The methodologies for deciding low-level saliency can be founded on natural models or computational models. Some methodologies consider saliency more than a few scales while other works on solitary scale. Here it is necessary to decide nearby balance of picture areas with its background support. As reported the HDCT-based method presented is one of the top six algorithms in salient region detection. Local-contrast-based models detect salient regions by detecting rarity of image features in a small local region. In Itti et.al paper they [3] proposed a saliency detection method which utilizes visual filters called "center-surround difference" to compute local color contrast. Harel et al. [4] suggested a graph-based visual saliency (GBVS) model which is based on the Markovian approach on an activation map. This model examines the dissimilarity of center-surround feature histograms. Goferman et al. [5] combined global and local contrast saliency to improve detection performance. Klein and Frintrop [6] utilized information theory and defined the saliency of an image using the Kullback-Leibler divergence (KLD). The KLD measures the center-surround difference to combine different image features to compute the saliency. Hou et al. [7] used the term "information divergence" which expresses the non-uniform distribution of the visual information in an image for saliency detection. Several methods estimated saliency in superpixel level instead of pixel-wise level to reduce the computational time. Jiang et al. [8] performed salient object segmentation with multiscale super pixel-based saliency and a closed boundary prior. Their approach iteratively updates both the saliency map and the shape prior under an energy minimization framework.

Perazzi et al. [9] decomposed an image into compact and perceptually homogeneous elements, and then considered the uniqueness and spatial distribution of these elements in the CIELab color to detect salient regions. Yan et al. [10] used a hierarchical model by computing contrast features at different scales of an image and fused them into a single saliency map using a graphical model. Zhu et al. [11] proposed a background measure that characterizes the spatial layout of image regions with a novel optimization framework. These models tend to give a higher saliency at around edges and texture areas that have high contrasts, where humans tend to focus on in an image. However, these models tend to catch only parts of an object. Also, they tend to give non-uniform weight to the same salient object when different features presented in the same salient object. Global-contrast-based models use color contrast with respect to the entire image to detect salient regions. These models can detect salient regions of an image uniformly with low computational complexity.

Achanta et al. [12] proposed a frequency-tuned approach to determine the center-surround contrast using the color and luminance in the frequency domain as features. Shen and Wu [13] divided an image into two parts—a low-rank matrix and sparse noise—where the former explains the background regions and the latter indicates the salient regions. Cheng et al. [14] proposed a Gaussian mixture model (GMM)-based abstract representation method that simultaneously evaluates the global contrast differences and spatial coherence to capture perceptually homogeneous elements and improve the salient region detection accuracy. Li et al. [15] showed that the unique refocusing capability of light fields can robustly handle challenging saliency detection problems such as similar foreground and background in a single image.

Borji and Itti [16] used complementary local and global patch-based dictionary learning for rarity-based saliency in different color spaces RGB and LAB and then combined them into the final saliency map for saliency detection. Jiang et al. [17] proposed a multilevel image segmentation method based on the supervised learning approach that performed a regional saliency regress or using regional descriptors to build a saliency map to find salient regions. These models are usually highly accurate and have a simple detection structure. However, they tend to require a lot of computational time. Therefore, superpixel-wise saliency detection is used to overcome the high computational complexity.

3. RELATED WORK

3.1 Efficient Salient Region Detection with Soft Image Abstraction

In this paper, we propose a novel soft image abstraction approach that captures large scale perceptually homogeneous elements, thus enabling effective estimation of global saliency cues. Unlike previous techniques that rely on super-pixels for image abstraction, we use histogram quantization to collect appearance samples for a global Gaussian Mixture Model (GMM) based decomposition. Components sharing the same spatial support are further grouped to provide a more compact and meaningful presentation. This soft abstraction avoids the hard decision boundaries of super pixels, allowing abstraction components with very large spatial support. This methods result is as shown in fig. 2.

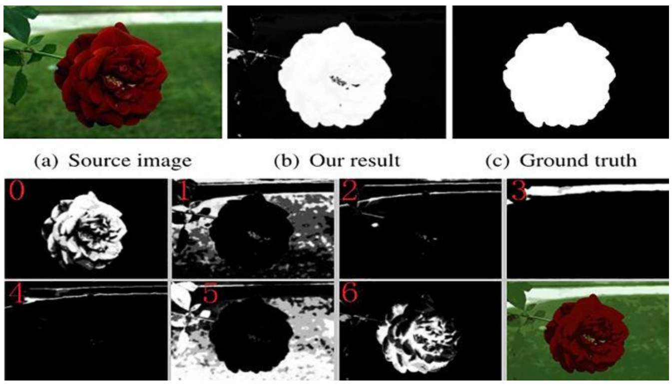


Fig -2: Soft image abstraction method to decompose an image into large scale.

This allows the subsequent global saliency cues to uniformly highlight entire salient object regions. Finally, we integrate the two global saliency cues, Global Uniqueness (GU) and Color Spatial Distribution (CSD), by automatically identifying which one is more likely to provide the correct identification of the salient region. We extensively evaluate our salient object region detection method on the largest publicly available dataset with 1000 images containing pixel accurate salient region annotations [18]. The evaluation results show that each of our individual measures (GU and CSD) significantly out performs existing 18 alternate approaches, and the final Global Cues (GC) saliency map reduces the mean absolute error by 25.2% compared to the previous best results, while requiring substantially less running times. In order to get an abstract global representation which effectively captures perceptually homogeneous elements, we cluster image colors and represent them using Gaussian Mixture Models (GMM). Each pixel color I_x is represented as a weighted combination of several GMM components, with its probability of belonging to a component c given by:

$$p(c\,|I_x) = \frac{\omega_c\, N(I_x\,|\mu_c\,,\Sigma_c)}{\Sigma_c\, \omega_c N(I_x|\mu_c\,,\Sigma_c)} \qquad (1)$$

where ω_c, μ_c, and Σ_c represent respectively the weight, mean color, and covariance matrix of the cth component. We use the GMM to decompose an image in to perceptually homogenous elements. These elements are structurally representative and abstract away unnecessary details fig. 2 shows an example of such decomposition. Notice that our GMM-based representation better captures large scale perceptually homogeneous elements than superpixel representations which can only capture local homogeneous elements.

3.2 Frequency-tuned Salient Region Detection

The true usefulness of a saliency map is determined by the application. In this paper we consider the use of saliency maps in salient object segmentation. To segment a salient object, we need to binaries the saliency map such that one's (white pixels) correspond to salient object pixels while zeros (black pixels) correspond to the background. We present comparisons with our method against the five methods mentioned above. In the first experiment, we use a fixed threshold to binaries the saliency maps.

In the second experiment, [21] we perform image-adaptive binarization of saliency maps. In order to obtain an objective comparison of segmentation results, we use a ground truth image database. We derived the database from the publicly available database used by Liu et al. This database provides bounding boxes drawn around salient regions by nine users. However, a bounding box-based ground truth is far from accurate, as also stated by Wang and Li . Thus, we created an accurate object-contour based ground truth database of 1000 images. To analyze the properties of the five saliency algorithms, we examine the spatial frequency content from the original image that is retained in computing the final saliency map. It understands that the range of spatial frequencies retained by our proposed algorithm is more appropriate than the algorithms used for comparison. For simplicity, the following analysis is given in one dimension and extensions to two dimensions are clarified when necessary.

In method IT, a Gaussian pyramid of 9 levels (level 0 is the original image) is built with successive Gaussian blurring and down sampling by 2 in each dimension. In the case of the luminance image, these results in a successive reduction of the spatial frequencies retained from the input image, its example is shown in fig. 3. Each smoothing operation approximately halves the normalized frequency spectrum of the image. At the end of 8 such smoothing operations, the frequencies retained from the spectrum of the original image at level 8 range within [0,π/256]. The technique computes differences of Gaussian-smoothed images from this pyramid, resizing them to size of level 4, which results in using frequency content from the original image in the range [π/256, π/16]. In this frequency range the DC (mean) component is removed along with approximately 99% ($(1-\frac{1}{16^2})\times 100$) of the high frequencies for a 2-D image. As such, the net information retained from the original image contains very few details and represents a very blurry version of the original image.

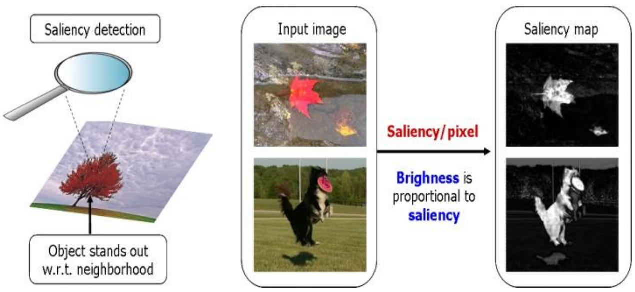


Fig -3: Frequency tuned salient region detection output.

3.3 Salient SLIC segmentation algorithm

Our approach [19] generates superpixels by clustering pixels based on their color similarity and proximity in the image plane. This is done in the five-dimensional [labxy] space, where [lab] is the pixel color vector in CIELAB color space, which is widely considered as perceptually uniform for small color distances, and xy is the pixel position. While the maximum possible distance between two colors in the CIELAB space (assuming RGB input images) is limited, the spatial distance in the xy plane depends on the image size. It is not possible to simply use the Euclidean distance in this 5D space without normalizing the spatial distances. In order to cluster pixels in this 5D space, we therefore introduce a new distance measure that considers superpixel size. Using it, we enforce color similarity as well as pixel proximity in this 5D space such that the expected cluster sizes and their spatial extent are approximately equal.

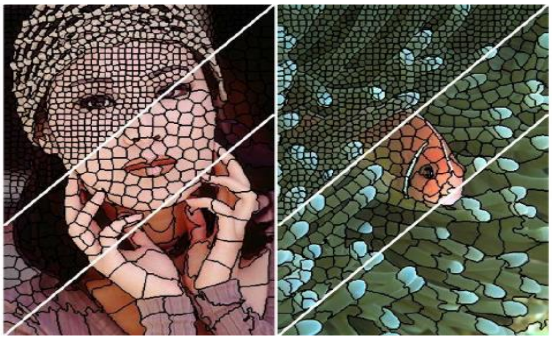


Fig -4: Using SLIC segmentation algorithm Superpixels generation.

Operating on superpixels instead of pixels can speed up existing pixel-based algorithms, and even improve results in some cases. For instance, certain graph-based algorithms can see a 2 to 3-fold speed increase using superpixels as shown in fig. 4. Of course, the superpixel generation itself should be fast for this to be practical. Below, we consider two typical vision tasks that benefit from using superpixels: object class recognition and medical image segmentation. In each case, superpixels have been shown to increase the performance of an existing algorithm while reducing computational cost. We show that SLIC superpixels outperform state-of-the-art superpixel methods on these tasks, but with a lower computational cost.

3.4 High Dimension Color Transform Algorithm

In this section, we describe our method to detect the initial location of salient regions in an image. Our method [20] is a learning-based method and it processes an image in super pixel level. The initial saliency tri-map consists of foreground candidate, background candidate, and unknown regions. A similar approach has already been used in a previous method, which demonstrated superiority and efficiency in their results. However, their algorithms require considerable computational time because their features' computational complexity is very large. In our work, we only use some of the most effective features that can be calculated rapidly, such as color contrast and location features. As our goal in this step is to "approximately" find the salient regions of an image, we found that the salient region could be found accurately using even a smaller number of features. By allowing for the classification of some ambiguous regions as unknown, we can further improve the accuracy of our initial saliency trimap.

The location feature is used because humans tend to focus more on objects that are located around the centre of an image. Next, we concatenate histogram features as this is one of the most effective measurements for the saliency feature. The histogram features of the ith super pixel DHi is measured using the chi-square distance between other super pixels' histograms. It is defined as

$$D_{Hi} = \sum_{j=1}^{N} \sum_{k=1}^{b} \frac{(h_{ik}+h_{ik})^2}{(h_{ik}+h_{ik})'} \qquad (2)$$

- Initial Saliency Trimap via Random Forest Classification

After we calculate the feature vectors for every superpixel, we use a classification algorithm to check whether each region is salient. In this study, we use the random forest classification because of its efficiency on large databases and its generalization ability. A random [3][6]forest is an ensemble method that operates by constructing multiple decision trees at training time and decides the class by examining each tree's leaf response value at test time.

$$D_{Li} = \sum_{j=1}^{N} \omega_{i,j}^{P} d(c_i, c_j) \qquad (3)$$

$$\omega_{i,j}^{P} = \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_P^2} \left\| \mathrm{P}_i - \mathrm{P}_j \right\|_2^2\right), \qquad (4)$$

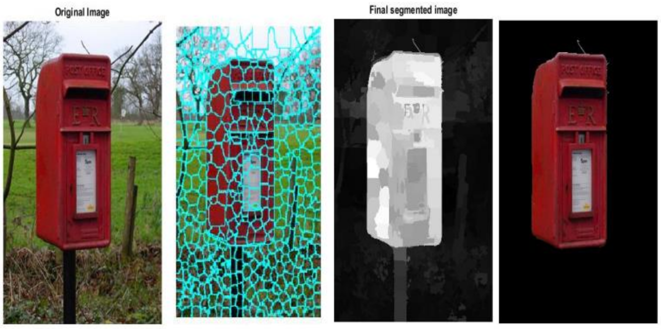
This method combines the bootstrap aggregating idea and random feature selection to minimize the generalization error. To train each tree, we sample the data with the replacement and train a decision tree with only a few features that are randomly selected. Typically, a few hundred to several thousand trees are used, as increasing the number of trees tends to decrease the variance of the model.

$$F_p = \frac{|\{F_C\} \cap \{F_{GT}\}|}{|\{F_C\}|} \qquad (5)$$

$$B_p = \frac{|\{B_C\} \cap \{B_{GT}\}|}{|\{B_C\}|} \qquad (6)$$

$$E_R = \frac{|(\{F_C\} \cap \{B_{GT}\}) \cup (\{B_C\} \cap \{F_{GT}\})|}{|\{I\}|} \qquad (7)$$

From the outputs of the random forest, we use a three-class classification to generate a trimap, instead of a binary classification, to detect highly reliable foreground/ background regions. Trimap has been commonly used in matting methods. In our work, we use the concept of trimap at the initial saliency estimation step. We set the relatively reliable regions of salient and non-salient regions to foreground or background respectively, and consider the ambiguous regions as unknown. Compared to the binary maps without unknown regions, we found that classifying ambiguous regions as unknown regions can help to obtain more reliable locations of salient regions. We decided whether each superpixel belongs to foreground candidate, background candidate, or unknown regions using the response value extracted. After performance above all levels in MATLAB software we got the resulting output shown in fig. 5 by Global saliency estimation via HDCT and local saliency estimation via regression. It understands that this method has better performance than previous results.



(a) (b) (c) (d)

Fig -5: Output of performed HDCT method. (a) Input. (b) Superpixel slicing. (c) Salient region detection with final segment. (d) Output of high dimension color transform method.

4. CONCLUSIONS

Object detection is necessity of picture handling process in image processing. In this paper we survey the current systems of salient object detection. Salient region detection is a challenging problem and an important topic in computer vision. The single visual cue based salient region detection methods have their own limitation. Implementation of HDCT method deals with salient region detection that estimates the foreground regions from a trimap using two different methods: global saliency estimation via HDCT and local saliency estimation via regression. The trimap-based robust estimation overcomes the limitations of inaccurate initial saliency classification. As a result, this method achieves good performance and is computationally efficient in comparison to the state-of-the art methods.

REFERENCES

[1] Z. Liu, R. Shi, L. Shen, Y. Xue, K. N. Ngan, and Z. Zhang, "Unsupervised salient object segmentation based on kernel density estimation and two-phase graph cut," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1275–1289, Aug. 2012.

[2] V. Navalpakkam and L. Itti, "An integrated model of top-down and bottom-up attention for optimizing detection speed," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2006, pp. 2049–2056.

[3] L. Itti, J. Braun, D. K. Lee, and C. Koch, "Attentional modulation of human pattern discrimination psychophysics reproduced by a quantitative model," in Proc. Conf. Adv. Neural Inf. Process. Syst. (NIPS), 1998, pp. 789–795.

[4] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Proc. Conf. Adv. Neural Inf. Process. Syst. (NIPS), 2006, pp. 545–552.

[5] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 2376–2383.

[6] D. A. Klein and S. Frintrop, "Center-surround divergence of feature statistics for salient object detection," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Nov. 2011, pp. 2214–2219.

[7] W. Hou, X. Gao, D. Tao, and X. Li, "Visual saliency detection using information divergence," Pattern Recognit., vol. 46, no. 10, pp. 2658–2669, Oct. 2013

[8] H. Jiang, J. Wang, Z. Yuan, T. Liu, and N. Zheng, "Automatic salient object segmentation based on context and shape prior," in Proc. Brit. Mach. Vis. Conf. (BMVC), 2011, pp. 110.1–110.12.

[9] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 733–740.

[10] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2013, pp. 1155–1162.

[11] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2014, pp. 2814–2821.

[12] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1597–1604

[13] X. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 853–860.

[14] M.-M. Cheng, J. Warrell, W.-Y. Lin, S. Zheng, V. Vineet, and N. Crook, "Efficient salient region detection with soft image abstraction," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2013, pp. 1529–1536.

[15] N. Li, J. Ye, Y. Ji, H. Ling, and J. Yu, "Saliency detection on light field," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2014, pp. 2806–2813

[16] A. Borji and L. Itti, "Exploiting local and global patch rarities for saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 478–485.

[17] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li, "Salient object detection: A discriminative regional feature integration approach," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2013, pp. 2083–2090.

[18] R. Achanta, S. Hemami, F. Estrada, and S. S¨usstrunk. Frequency-tuned salient region detection. In IEEE CVPR, pages 1597–1604, 2009.

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[20] J. Kim, D. Han, Y.-W. Tai, and J. Kim, "Salient region detection via high-dimensional color transform," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2014, pp. 883–890.

[21] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1597–1604. 22 IEEE transactions on image processing, vol. 25, no. 1, january 2016.