IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Fusion of Saliency Based Co-Saliency Detection

M.Sreenavya[1], Chandra Mohan Reddy Sivappagari[2],

[1, 2]Department of Electronics and Communication Engineering, JNTUA College of Engineering, Pulivendula, Andhra Pradesh, India

Corresponding Author: manchurunavya@gmail.com, Mob: +91-7036909113

Available online at: www.ijcseonline.org

Accepted: 16/Jul/2018, Published: 31/July/2018

Abstract - Co-saliency is utilized for exploring common saliency present within numerous pictures or images and is an area of research that is still under explored. This not only deals with the visual cues present within the images but also covers the cues that are outside the image and hence deals with the shortcoming present within saliency detection of a single-image. It depends upon the visual cues that are already discovered or explored and varies from place to place. In order to address this concern, this paper aims to propose a technique that can be helpful in detecting the co-salient objects on map fusion and are region-wise saliency. This technique takes into account the intra image appearance, its correspondence with the features outside the image, the spatial features or factors and aims at the detection of salience with the help of a saliency that is locally adaptive map fusion through dealing with the problem within the map in relation to energy optimization. This technique or method will be accessed on the basis of a standard dataset that is taken as a benchmark and is compared with other techniques and methods that are available.

Keywords — Co-saliency detection, graph-based optimization, energy minimization, locally adaptive fusion.

I.INTRODUCTION

Saliency detection is a preferential method or technique utilized for distributing resources that are computational in nature. The present algorithm formulas for saliency help in detecting objects that are salient within each image. In the present times, correspondence for multiple images is dependent upon a small set or group of images and has become a challenging issue. Co-saliency detection is described as a process utilized for exploring a unique item or object within a set of analogous images. However the condition of images to be shot in a similar burst of shots, limit its implementation.

Saliency detection's main purpose is to recognize the salient pixels within an image without any supervision [1]. It is a very significant area of research in the field of image processing, as it helps in the automation of numerous applications including image segmentation [2] and video compression just to name a few [3]. Performances of saliency detection within these single-images has certain limitations and are unsubstantiated in nature, particularly the images that are complex. This detection process has been introduced in order to overcome the issues present within the saliency detection of a single-image and aims to explore as well as locate the salient object or items that are common in these images.

The literature that is already available in relation to the approaches present for detection can be divided into two types; inter and intra image evidences or proofs. The inter-image are attained by detecting the nature of communications between the pair of images whereas the latter is drawn on the grounds of spatial cues as well as the appearance contrast. The above-mentioned approaches blend saliency maps in a sequential order, which is the weight that is given to the saliency map within the image as a whole. These methods however do not pay heed towards the fact that these saliency maps are dependent upon regions most often.

Ground truth utilizes a group of measurements that are more precise and accurate in nature than the measurements deduced from other systems which are employed for testing and measurement purposes. These are employed to numerous areas such as machine learning, satellite imagery, etc. In addition, the ground truth is employed on an extensive level for image processing including edge detection, image segmentation, etc.

This paper proposes will be proposing a technique or method that takes into account, the inter and intra image proofs and helps in initiating saliency maps that are drawn from regional-based fusion. This method is helpful in decreasing the unwanted effects in the form of misses and false or wrong alarms, and lead towards the generation of saliency maps that are of optimum quality.

In this paper, Section I contains the introduction of related work system , Section II contain the related work of previous papers, Section III contain the methodology, Section

IV contain the results and discussion, section V contains the conclusion and future scope.

II. RELATED WORK

The quantity of literature available on this theme is extensive in nature [4, 5] and deals with the detection of salient objects effectively. An approach that deals with the process of eye fixation prediction is stimulated by the basic visual system of humans. An attention mechanism that is visual in nature is deduced from an architecture that is neuronal based and helps in stimulating behaviour. Multiscale image features are joined together to generate a saliency map that is topographical saliency map. An active neural network than looks for locations that have decreasing or diminishing degrees of saliency. The complex problem is broken down into pieces and make selections through a computation based process that takes into account the details of each location.

In this each characteristic is calculated with the help of a set of linear center that are surrounded by receptive visual fields. The typical visual neurons are susceptible within areas that have small visual space, whereas stimuli which is present in a broad or large area leads to a neuronal response in relation to the center or visual space. This approach fails in detecting the borders of the object however is capable to measure natural scenes that are complex in nature [4].

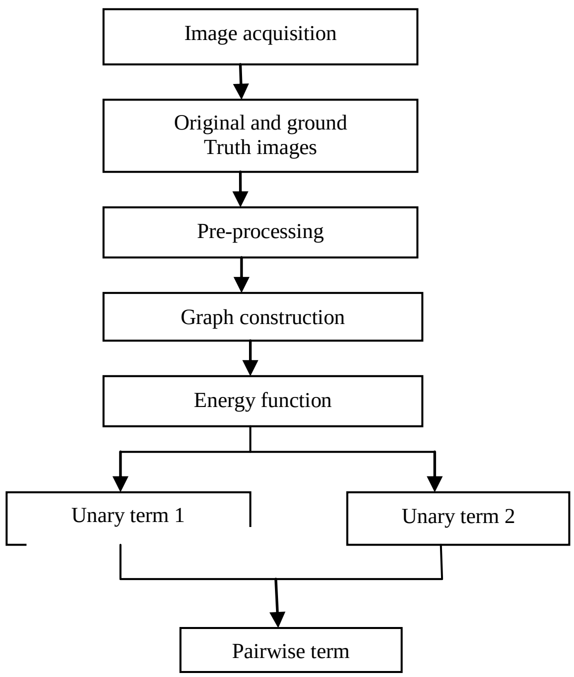
The residual on the log-frequency is a method for detecting maps saliency. Though this technique can be computed or calculated, it most of the times determines the boundaries of the objects rather than determining salient areas or regions completely. The system that works with the aim of reducing the redundancy in the visual features should know about the similarities that are statistical in nature in the input image. Hence in numerous log spectra that can be used to detect the similarities between shapes, the information that is most vital is the one that jumps outside the smooth curves. The similarities between the statistical singularities in the spectrum can lead to regions that are anomalous in nature within image, and are points where proto-objects pop up. [4, 5] These two methods include process related to image resizing, which possibly leads to loss within the content frequency.

The salient object detection, [6] developed a complete method through which salient regions were highlighted in a uniform manner and precise boundaries were drawn with it. However this method category ignored the spatial features of the items within an image and predicted the background as salient regions or areas. Shen and Wu [7] also combined high-level features like the semantic prior or the center while detecting items or objects. Another model [8] amended Achanta et al.'s model by taking into account the spatial features and appearance contrast while detecting salience. Apart from the low-level features, Yang et al. [9] utilized background to infer boundaries of objects and foreground to rank the degrees of saliency within the super pixels. In [9], Zhu et al. [10] suggested a new method for generating the background prior.

Co-saliency detection is proposed to improvise the performance levels. The combined visual cues that are derived from the images assist in the removal of background as well as foreground. For example, Li and Ngan [11] used the SimRank algorithm on a co-multilayer super pixel tree, and identified the similarities in the textures as well as color between the pixels of each image. Meng et al. [12] enhanced the SimRank matching technique by taking into account the limitation or constraints that are geometric in nature. Fu et al. [13] suggested a process that was cluster-based in order to gain an insight of the inter-image correspondence. Cao et al. [14, 15] applied a low-level limitation on the saliency regions or areas on the map proposals of multiple salience and determined the fusion weight for each one. This approach takes into account the fact that saliency maps are region-based and hence seek weights with the help of fusion process in a region-based sequential order, hence leading to encouraging and positive outcomes.

III. METHODOLOGY

Here the pair of images I_1 and I_2 for co-saliency detection, Apply M existing co-saliency detection algorithms, for instance the research presented in previous papers [4, 5, 6, and 11] gets M saliency maps for each image. For fusion that is local-based or region-based, images I_1 and I_2 are individually divided into N_1 and N_2 super pixels, which work as region-based fusion domain. This technique aims to compute the weight vector for a region-wise fusion domain. This approach explores a weight vector $y_i = \begin{bmatrix} y_{i,1} & y_{i,2} & \cdots & y_{i,M} \end{bmatrix}^T \in R^M$ or each of the super pixel i, where $i \in \{1,2,\ldots,N_1+N_2\}$.The co-saliency detection is shaped by fusion that is super pixel-based M saliency maps. This process represents region-based fusion, across a graph in order to reduce energy optimization. The following graph represents the image pre-processing and its construction. The projected energy function and its optimization are defined later.



d saliency map

I proposed saliency map. F l ground truth. The ground

Pre-p

Here th C [16] algorithm is incorporated for drawing out t per pixels, because it efficiently maintains the intr structure and removes information that is pointless. Set the numbers of super pixels to $N_1 = N_2 = 200$ in this work.

Each super pixel helps in extracting two types of visual characteristics or features including color and texture. As mentioned below the method of RGB, L*a*b* and YCbCr color spaces are utilized in a combination to represent the characteristics of color and are adjusted within a range of 0 to 1. To generate a color descriptor for the region or area, firstly demonstrate a pixel as a 9-dimensional (9-D) color vector by joining parts or components of RGB, L*a*b* and YCbCr color spaces. Then all the pixels that are present in the form of groups or pairs are quantified into clusters with the help of k-means clustering algorithm, where each cluster center is labelled as a code word.

For texture features, Gabor filter is utilized that includes eight directions or orientations, three scales and two phase offsets are drawn for each pixel. The texture characteristics of the pixel are encoded as a 100-dimensional histogram.

Let p_i and q_i denote the color and texture of each super pixel i respectively and the correspondences amongst the super pixel i and superpixel *j* is defined as

$$A(i,j) = \exp\left(-\frac{d(p_i, p_j)}{\sigma_c} - \gamma \frac{d(q_i, q_j)}{\sigma_c}\right), \qquad (1)$$

In the Graph construction segmenting super pixels.

Energy Function

This intends to measure the optimal weights $Y = [y_1 \;\; y_2 \;\; \ldots y_N] \in R^{M\times N}$ Where M is the amount of saliency maps, and *N* is the total amount of super pixels of I_1 and I_2, for super pixel-wise map fusion reducing the amount of proposed or suggested energy function

$$\min_Y \lambda_1 \sum_{v_i \in v} U(y_i) + \lambda_2 \sum_{v_i \in v} V(y_i) + \lambda_3 \sum_{e_{i,j} \in \varepsilon} B(y_i, y_j) + \|Y\|_2^2 \qquad (2)$$

Four terms have been familiarized in Eq. (2). The first two are unary terms, $U(y_i)$ and $V(y_i)$, individually influence both the inter-image and intra-image evidences, in order to estimate the power of each saliency map on super pixel i. The pairwise term $B(y_i, y_j)$ helps in drawing the weights on super pixel pairs connected in the graph *g*. The last term $\|Y\|_2^2$ is included for regularization.

Unary Term $U(y_i)$

Here calculate the weight of saliency map. Let $X_{i,m} \in R^{256}$ be a 256-dimensional histogram that represents the intensity distribution of the saliency values of the map m on these *n* super pixels. By stacking the M different vectors for all saliency maps, $X_i = [X_{i,1} \;\; X_{i,2} \;\; \cdots \;\; X_{i,M}] \in R^{256\times M}$.

Unary Term $V(y_i)$

This term is planned in order to diminish saliency detection that is false in nature through discovering the evidences which are present within inter-image communications.

Pairwise Term $B(y_i, y_j)$

Here this term represents joining the two unary terms. The smoothness within the distribution of weights Y are connected between super pixels.

IV. RESULTS AND DISCUSSION

This approach is accessed in this section through image Pair dataset [11], which includes about 105 pair of images which are labelled manually with ground truth. Here the investigational arrangement for accessing the approach is measured with the help of a Precision-Recall (PR) curve, which is acquired with the help of a saliency threshold that varies. PR curve incline in the favor of this method and highlight that it is productive in detecting salience and precisely locates regions that have no salience. Therefore, Receiver Operating Characteristics (ROC) curve and Mean Absolute Error (MAE) are grounded on the truth and also help in accessing the performance levels.

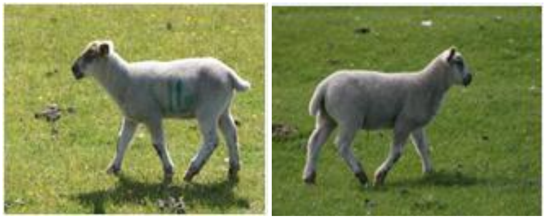


Fig 2(a):Input image pair

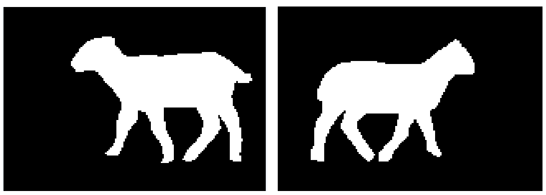


Fig 2(b):Ground truths

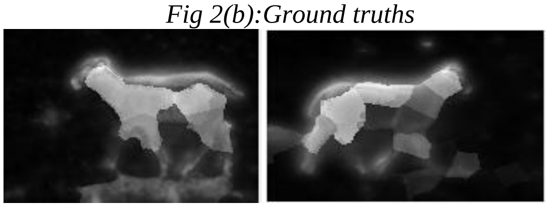


Fig 2(c):Saliency maps

The above fig 2(a) and 2(b)shows one example image pair and ground truth. Here ground truth used for comparison. Fig 2(a) is a image pair same image with present different background here take. This approach using the region wise saliency map algorithm finally get the image shows fig2(c).

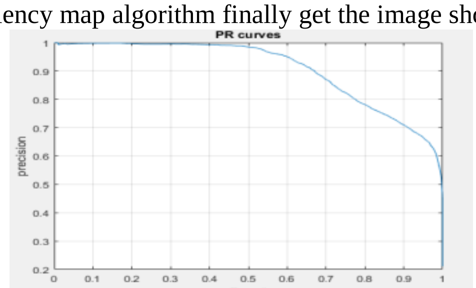


Fig 2(d): PR Curve

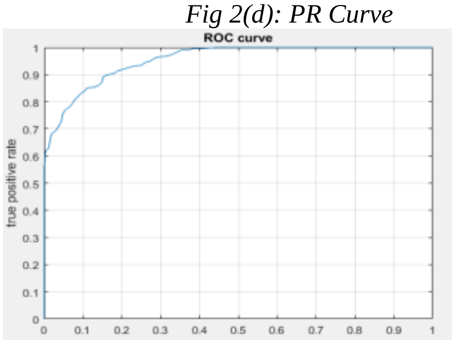


Fig 2(e): ROC Curve

The above fig 3(a) and 3(b) shows the PR curve and ROC curve. Here the higher value is the better value. The remaining image pairs follows the above process.

Fig 3(a):Input image pair

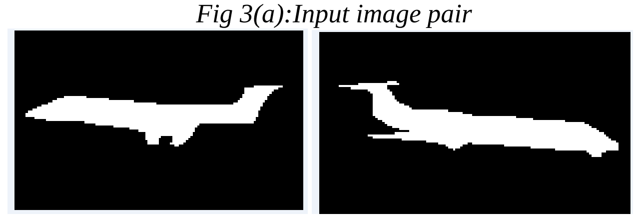


Fig 3(b):Ground truths

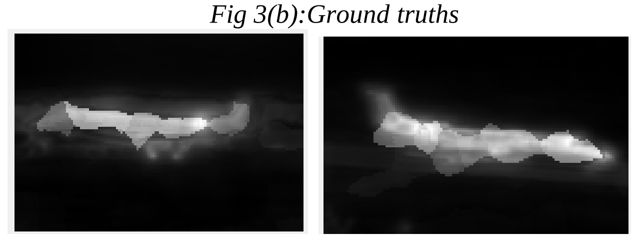


Fig 3(c): Saliency maps

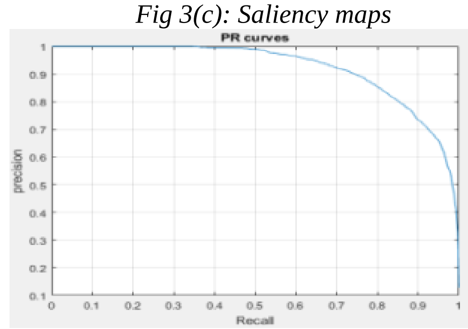


Fig 3(d): PR Curve

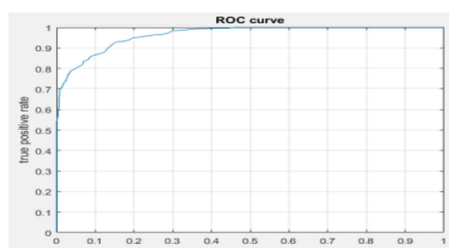


Fig 3(e): ROC Curve

Fig 4(a):Input image pair

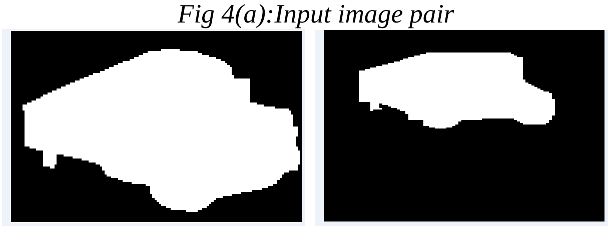


Fig 4(b):Ground truths

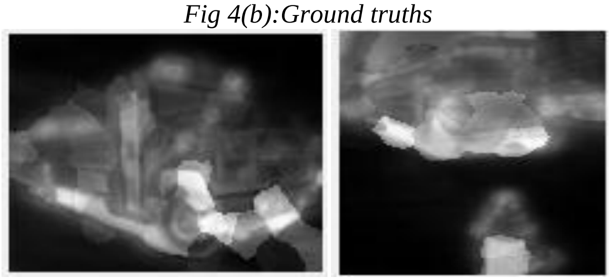


Fig 4(c): Saliency maps

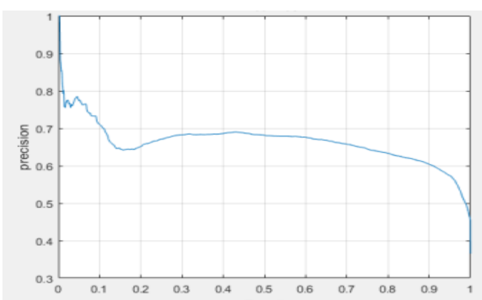


Fig 4(d):PR Curve

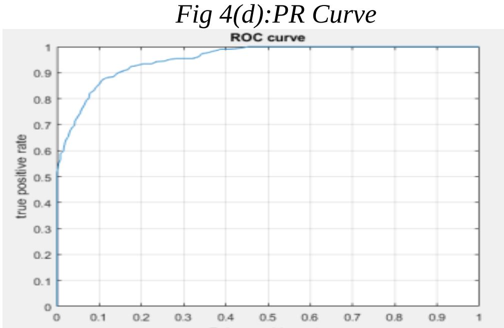


Fig 4(e): ROC Curve

Table 1: The performance of 1) AUC of PR, 2) AUC of ROC, and 3) MAE.

Parameter	Fig 2	Fig 3	Fig 4
PR AUC	1	1	0.9
ROC AUC	1	1	1
MAE	0.14	0.083	0.13

The PR curve and the ROC curve from this method also report the area under curve (AUC) of the PR curve, the AUC of ROC curve, and MAE of these approaches as mentioned in Table 1. This approach generates saliency maps through the fusion process that is regional-based or local-based. As shown this method consistently achieves better performance.

To get an understanding of the quantitative result, Fig. 1 highlights the saliency map detection of each pair of image. It uses inter-image evidences for the saliency proposal and also includes the regions that have false saliency. In the meantime, the saliency proposal utilizes these evidences and this method through the fusion process that is region-based fusion and fulfils the ground truths which are closest to the saliency map. Moreover, the saliency map with the assistance of this approach become sharper in detecting the salience with precision and confidence.

V. CONCLUSION AND FUTURE SCOPE

The above mentioned approach for detecting salience uses a saliency map fusion that is adapted locally. It helps in generating saliency maps that are region-based and generates optimum quality saliency maps both quantitatively and perceptually. In forthcoming, this plan can help in accessing this approach and setting benchmark datasets and work together with tasks that are related to one another such as co-segmentation.

REFERENCES

[1]. Chung-Chi Tsai1, Xiaoning Qian and Yen-Yu Lin,” Image co-saliency detection via locally adaptive saliency map fusion”, ICASSP 2017.

[2]. Kai-Yueh Chang, Tyng-Luh Liu, and Shang-Hong Lai,"From co-saliency to co-segmentation: An efficient and fully unsupervised energy minimization model," in proc.conf.Computer Vision and Pattern Recognization, 2011.

[3]. Zhicheng Li, Shiyin Qin, and Laurent Itti, "Visual attention guided bit allocation in video compression," j.Image and Vision Computing, 2011.

[4]. Laurent Itti, Christof Koch, and Ernst Niebur, "A model of saliency-based visual attention for rapid scene analysis,"IEEE Trans.on Pattern Analysis and Machine Intelligence, 1998.

[5]. Xiaodi Hou and Liqing Zhang, "Saliency detection: A spectral residual approach," in Proc.Conf. Computer Vision and Pattern Recognization, 2007.

[6]. Ravi Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk, "Frequency-tuned salient region detection,"in Proc.Conf. Vision and Pattern Recognization, 2009.

[7]. Xiaohui Shen and Ying Wu, "A unified approach to salient object detection via low rank matrix recovery,"in Proc. Conf. Vision and Pattern Recognization, 2012.

[8]. Federico Perazzi, Philipp KrahenbUhl, Yael Pritch, and alexander Hornung, "Saliency filters: Contrast based filtering for salient region detection," in Proc.Conf. Vision and Pattern Recognization, 2012.

[9]. Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang, "Saliency detection via graph based manifold ranking," in Proc.Conf. Vision and Pattern Recognization, 2013.

[10]. Wangjiang Zhu, Shuang Liang, Yichen Wei, and Jian Sun, "Saliency optimization from robust background detection," in Proc.Conf. Vision and Pattern Recognization, 2014.

[11]. Hongliang Li and King Ngi Ngan, "A co-saliency model of image pairs,"IEEE Trans. On Image Processing 2011.

[12]. Fanman Meng, Hongliang Li, and Guanghui Liu, "A new co-saliency model via pairwise constraint graph matching," in IEEE Int'l Symposium, Intelligent Signal Processing and Communications System, 2012.

[13]. Huazhu Fu, Xiaochun Cao, and Zhuowen Tu, "Cluster based co-saliency detection," IEEE Trans. On Image Processing, 2013.

[14]. Xiaochun Cao, Zhiqiang Tao, Bao Zhang, Huazhu Fu, and Wei Feng, "Self-adaptively weighted co-saliency detection via rank constraint," IEEE Trans. On Image Processing, 2014.

[15]. Xiaochun Cao, Zhiqiang Tao, Bao Zhang, Huazhu Fu, and Xuewei Li, "Saliency map fusion based on rank-one constraint," in Proc. Int'l Conf Multimedia and Expo, 2013.

[16]. Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk, "SLIC superpixels," Tech. Rep., 2010.

AUTHOR'S PROFILE

M. Sree Navya studied B.TECH from Vemu institute of technology, Chittoor, Andhra Pradesh. And now pursuing Master of Technology in Digital Electronics and Communication Systems from JNTUA College of Engineering, Pulivendula, Andhra Pradesh.

Dr. Chandra Mohan Reddy Sivappagari is currently working as Associate Professor of ECE, JNTUA College of Engineering Pulivendula, Andhra Pradesh, India since 2006. He is a member of IEEE, life member of IE (I), ISTE and IAENG. He has published more than 40 research papers in reputed both in international and national journals and conferences. His main research work focuses on Signal & Image Processing and Internet of Things. He has 17 years of teaching experience and 8 years of research experience.

