International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 3, Special Issue 3 (September '2017)
Proceeding of - International Conference on Recent Trends in Computational Life Science and Information Technology,
Conference held at Tiruppur Kumaran College for Women,Tirupur,Tamilnadu, India.

REGION BASED IMAGE RETRIEVAL USING FUZZY ORIENTED SALIENCE SIMILARITY ESTIMATION METHOD

M.Shanthakumar,
Assistant Professor,
Department of Computer Science,
Gobi Arts and Science College,
Erode, Tamilandu, India.

S.Janarthanam,
Assistant Professor,
Department of Computer Science,
Gobi Arts and Science College,
Erode, Tamilandu, India.

S.Sukumaran,
Associate Professor,
Department of Computer Science,
Erode Arts and Science College,
Erode, Tamilandu, India.

Abstract: The development of detectable points is important in image processing the effective fuzzy oriented saliency detection make advances in image contents. Usually the take the role normal are identified by bilateral filters and retaining the local features. Fuzzy oriented method with patterns between the center pixel and its surrounding neighbours in two dimensional local region proposed robust fuzzy salient region extraction algorithm (FSE) encodes the spatial relation between any pair of neighbours in a local region along the directions for the center pixel in an image. The experiments carried out for proving the worth of proposed method on two different types of benchmark databases and the new metrics Directional based Gaussian weight (DGW) and Relative Normal Distance (RND) has been proposed in this paper. The comparative studies based on absolute data from two publicly accessible databases show that the proposed method usually outperforms both qualitatively and quantitatively for saliency estimation analysis and understanding. It can be further extended with implementation of new retrieval techniques with different parameters which improves more retrieval efficiency and performance integrity.

Keywords:Bilateralfilter;Gaussianfilter;Imagecontent;Spatial- feature; Saliency.

I. INTRODUCTION

This paper exploring the combination of features describe an image with respect to its visual properties, visual content and specifically focusing on content-based image retrieval (CBIR) tasks on regions The astute attempts to factorize concede components trifling by extracting great image features such as colour, texture and shapes that are calculated over the entire image. The consummate appropriately of extracting global features is the low cost of the single feature space computations. In addition, a global vector representation is a very effective strategy for certain retrieval tasks [1]. The sematic cleft to retain some degree of human interventions for preprocessing and recognition ability effectively for sematic retrieval, Keep a pursue the bilateral filter finds the local features and prevents the large differences in the normal vectors from the region of interest based on the statistics of the Euclidean distance between the normal of neighbors[2]. For invariance difference transformations to do the performance comparison of Haar Transform and Discrete Cosine Transform are used for the similarity measures. The pioneering results based on real data estimated the features more expressive in surface details [3,4]. The retrieved images are ranked according to the relevance between the query image and images in the database in proportion to a similarity measure calculated from the features. The remainder of this paper organized as follows. The region based and proposed saliency similarity estimation methods are explained in section II, III and IV respectively. Then the experimental results of the detected saliency to guide the existing techniques for different tasks are demonstrated in section V. Finally the conclusion in section VI.

II.RELATED WORK

The Lu and Burkhardt introduced an integrated method to color image retrieval by combining vector quantization method and discrete cosine transform. Ashraf et al. proposed image retrieval by using a combination of color features and bandlet transformation [5]. The transformation-based representation was selected to extract the salient objects from the images. Fang et al. propose a fuzzy logic approach for temporal segmentation where color histogram intersection, motion compensation, texture change and edge variances are integrated for cut detection[6].

A viewpoint determination system indigent upon the figuring recommended by Eesa, Orman, and Brifcani select sub-set starting with asserting Characteristics. The crash from claiming the individuals picked Characteristics used a decision tree classifier. The individuals prudent around fulfill higher rates something like precision concerning delineation for the each particular case picked [7,8]. The individuals on the most elevated necessity ahead related meets desires the individuals outstanding territories beginning with a picture a chance to be suggested looking into utilize the general qualities of the including region of the pixels. Those settle to saliency reasoned clinched alongside light for the individuals' voluminous correlations of the enhanced picture. Saliency will make a relative concept, judged not only against neighborhood neighbors, and also overall general state [9]. Focuses geometric properties starting with guaranteeing an object

described inevitably examining the neighboring concentrates. The global geometry might be the geometric properties to describe every single particular case in the dataset should portrayal.

III.PROPOSED WORK

The proposed region based saliency estimation consists the following phases as main steps bilateral normal filtering, fuzzy oriented vertex based saliency estimation description and directional oriented region based saliency estimation and understanding. This is the modified, extended work [10] and it has been improved in the sense that the bilateral filter has been proposed to smooth normal surfaces, the mean curvature feature has been replaced by the relative normal distance and validate more extensively using the real experiment data. The region-based directional saliency estimation method is proposed in this paper in order to refine the vertex-based saliency, through considering the votes of different surface regions for estimating their relative contrast and saliency values instead[10]. The final, fuzzy oriented vertex based saliency is estimated as a ratio between the region-based saliency and the vertex based saliency.

Directionality of an image is measures by the frequency distribution of oriented local edges against their directional angles. The histogram of quantized direction values are constructed by counting number of the edge pixels with the corresponding directional angels. Direction of edges that co-occurred in the pairs of pixels separated by a distance along the edge direction in every pixel using line-likeness regularity [11]. To validate the proposed method the three popular benchmark databases are used for the subjective nature of ground truth, the detected saliency is mainly presented for visual inspection. The usefulness of the detected saliency used to guide the existing techniques for directional oriented shape analysis and understanding tasks as mesh simplification, interest point detection, and overlapping point registration [12].

A. *Normalised Bilateral Filter*

Bilateral filter is the combination of two Gaussian filters: a spatial one, and a range one. The spatial Gaussian acts like the normal Gaussian filter in assigning larger weights to nearby pixels and smaller weights to distant pixels [13,14]. The main idea is to give more weight to the pixels that are closer in the spatial, but also in the range domain, to the current central pixel. The larger the distance in the spatial domain and the difference in the intensity domain, the less impact a neighboring pixel will impose on the pixel of interest [15]. The weight of each component is computed using a Gaussian function. It has been extended to filter the meshes due to its nonlinear, feature-preserving characteristics. The bilateral filter takes a weighted sum of the pixels in a local neighborhood; the weights depend on both the spatial distance and the intensity distance [16]. In this way, edges are preserved well while noise is averaged out. Mathematically, at a pixel location x, the output of a bilateral filter is calculated as follows,

$$\bar{I}(x)=\frac{1}{C}\sum_{y\in N(x)} e^{\frac{-\|y-x\|^2}{2\sigma_d^2}} e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}} I(y) \qquad (1)$$

Where σ_d and σ_r are parameter controlling the fall off of weight in spatial and intensity domains [17]. $N(x)$ is a spatial neighborhood of center pixel $I(x)$, I(y) is the any other pixel other than the center pixel in the given window and C is the normalization constant

$$C=\sum_{y\in N(x)} e^{\frac{-\|y-x\|^2}{2\sigma_d^2}} e^{\frac{-|I(y)-I(x)|}{2\sigma_r^2}} \qquad (2)$$

The spatial linear filters used to smooth noise from an image are the box (averaging) filter and the spatial Gaussian filter. These filters take the weighted average of the pixels in a neighborhood. An averaging filter has filter coefficients are equal, is called a box function The processed image is g(x, y) and M×N is the number of pixels being averaged on the input image f(x,y) as,

$$g(x,y)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{1}{M\times N} f(x_i,y_j)\right) \qquad (3)$$

Filters can also perform a weighted average with varying coefficients. A common averaging filter with this quality is the spatial Gaussian filter. Since the absolute distance between normal vectors tends to capture the coarse and large variation and ignore the fine and small variation in the surface normal, a new bilateral normal filtering method is proposed in this section. Given a unit normal vector n_i and the centroid c_i, the proposed bilateral filter normal $\bar{n}$ is defined as,

$$\bar{n}=\frac{\sum_{j=F_i} W_c(\|c_i-c_j\|)W_s(n_i,n_j)n_j}{\sum_{j=F_i} W_c(\|c_i-c_j\|)W_s(n_i,n_j)} \qquad (4)$$

Where $w_c=\exp\left\{-\frac{\|c_i-c_j\|^2}{2\sigma_c^2}\right\}$ is the spatial weight function and $w_s=\exp\left\{-\frac{\|n_i-n_j\|^2}{2\sigma_s^2}\right\}$ is the standard Gaussian function in terms of relative normal distance RND (n_i,n_j) and σ_c and σ_s are the two standard deviations of normal units and w be the spatial weight on the region under normal value but the over running weight recognized as pure value equal to 1. The relative normal distance RND(n_i,n_j) between n_i and n_j is defined as:

$$RND(n_i,n_j)=\frac{\|n_i-n_j\|}{w\times avg_{k\subset F_i}(\|n_i-n_k\|)} \qquad (5)$$

Where the average Euclidean distance between n_i and other neighboring normals n_k. The relative distance is not uniform. For two sets of points with a similar neighboring relationship between the corresponding points are differ but the relative distances are similar in general manner.

B. *Directional Gaussian Weight*

Frequency-tuned approach using a Difference of Gaussian filter helps uniformly highlight whole salient regions with well-defined boundaries. By adaptive feature refinement method by using low-level features are combined and

successively refined. The lightness distance, Color distance and edge strength are combined with weights are selected [17].

Let X denotes the m×n pixel input image, divide the block into equivalent region overlapped each other. Let *x* denote a block of X, and $f_i(x)$ be the feature value measure and $\forall f_i(x), x \in X$. Due to experiments the vision sensitive pattern contributes a little variations within the accumulate and saliency computation. The lightness and color difference of each block *x* is denoted by $f_1(x)$ and $f_2(x)$ are the computed Euclidean distance between the averages shown as, $f_1(x) = |\bar{L}^*(x) - \bar{L}^*(X)|$ and $f_2(x) = \sqrt{(\bar{a}^*(x) - \bar{a}^*(X))^2 + (\bar{b}^*(x) - \bar{b}^*(X))^2}$

Where $\bar{L}^*, \bar{a}^*, \bar{b}^*$ are the averages measured in CIELAB color space. In this method the random walk identifies the salient region determined with the help of frequency of visit of each node. Fully connected digraph representation of the edge between two vertices is proportional to their dissimilarity, as well as their closeness in the spatial domain [18,19]. Regions closer to the center of the image usually have a better chance of containing salient objects [20]. Weighting function was used to aggressively assign more weight toward the center of estimated location of the single main subject. However, for images containing multiple salient objects. Gaussian weighting function that serves primarily to suppress regions near the borders and corners of the image. The weights factors calculated by Gaussian bell $w(x,y) = e^{-(x^2+y^2/2r^2)}$, and (x, y) are arbitrary pixel positions, filter radius r is in statistics the standard deviation sigma relative to the filter box[21]. For original image the neighborhood distribution of image in local region are calculated as

$$I_x(x,\sigma_D) = \frac{\partial}{\partial x} g(\sigma_D) \times I(x) \qquad (6)$$

The local image derivatives are computed with Gaussian kernels of scale σ_D Gaussian kernel value of average in the neighborhood of the point by smoothing with a eigenvalues of this matrix represent the principal signal changes in two orthogonal directions in a neighborhood around the point image signal varies significantly in both directions, or in other words, for which both eigenvalues are large [22]. The gradient distribution of the smoothened image as,

$$g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-iw(x,y)} \qquad (7)$$

The sub-pixel precision can be achieved through quadratic approximation of the corner function in the neighborhood of a local maximum. Here, we employ a broad Gaussian weighting function that serves primarily to suppress regions near the borders and corners of the image[21,23]. Specifically the features modified as, $\bar{f}_i = f_i(x) \times f_c(x)$, where $f_c(x)$ be the Gaussian map function at the location(x,y) as,

$$f_c(x,y) = e^{-\left(\frac{(x-m/2)^2}{2\sigma_D^2} + \frac{(y-n/2)^2}{2\sigma_D^2}\right)} \qquad (8)$$

This property is attributable to the large values of the standard deviation σ_x and σ_y. Gaussian center-weighted function to normalize each feature map as follows

$$\bar{f}_i^{t+1} = \alpha \sum_{x_j \in N(x_i)} w_{ij} f_j^t + (1-\alpha) y_i \qquad (9)$$

Where f^t the classifying function of iteration and α is specifies the relative contribution of neighbor on the initial normal value [24]. The normalized saliency map of **X** denoted as $\mathbf{M_X}$, is the computed weight of all feature maps is in the range [0, 1].

$$M_x = \frac{\sum_i w_i \bar{f}_i^t}{\sum_i w_i} \qquad (10)$$

C. *Proposed Approach*

In order to develop a robust fuzzy salient region extraction algorithm (FSE) one needs to identify the problems incorporated with the available techniques. The fuzzy set theoretic based approach may be considered a good choice for handling uncertainties arising at different stages of processing and analysis of a CBIR system. A feature evaluation mechanism is provided to enhance the accuracy of the system further. The user marks the relevant images within the retrieved set. The individual feature weights are updated with a measure namely fuzzy feature evaluation index (FEI) computed from 'intra set ambiguity' and the 'interest ambiguity' of the relevant and irrelevant set of images.

Let $C_1, C_2, \ldots\ldots\ldots C_m$ be the m pattern classes in an N dimensional $(x_1, x_2, \ldots\ldots\ldots x_N)$ feature space class C_j contains n_j number of samples. But saliency originates from visual uniqueness, is often attributed to images in the fast visual attention focus on bottom up data driven saliency detection using image contrast. Saliency of a region depends mainly on its contrast to the nearby regions, while contrasts to distant regions are less significant.Let $I(x,y)$ be the image intensity at (x, y), where x=1, 2,…...W and y=1, 2,…….H, with W and H are the image width and height respectively. The Image representation as a set of pixels and is defined as

$$I = \{p_{xy} \mid x, y \in N \wedge x = 1,2.....W, \wedge y = 1,2,.....H\} \quad and \quad p_{xy} = I(x,y) \qquad (11)$$

Where p_{xy} is the pixel's value at location (*x*, *y*) and N is the set of Natural Numbers. Saliency maps should be fast and easy to generate to allow processing of large image collections, and facilitate efficient image classification and retrieval[43]. To evaluate saliency of an image region using its contrast with respect to the entire image. The saliency value of a pixel I_k in the image I is defined as,

$$S(I_k) = \sum_{\forall I_i \in I} D(I_k, I_i) \qquad (12)$$

Where D (I_k, I_i) is the color distance metric between pixels I_k and I_i. Since directly introducing spatial relationships when computing pixel-level contrast is computationally expensive, so as to integrate spatial relationships into region-level contrast computation (RC) method is used. The weights are set according to the spatial distances with farther regions being assigned smaller weights. Incorporate spatial information by introducing a spatial weighting term in Equation 18 to increase the effects of closer regions and decrease the effects of farther regions. Specifically, for any region r_k, the spatially weighted region contrast based saliency is defined as:

$$S(r_k) = \sum_{r_k \neq r_i} e^{(D_s(r_k, r_i)/\sigma_s^2) w(r_i) D_r(r_k, r_i)} \qquad (13)$$

Where $D_s(r_k, r_i)$ is the spatial distance between regions r_k and r_i, and σ_s controls the strength of spatial weighting. Larger values of σ_s reduce the effect of spatial weighting so that contrast to farther regions would contribute more to the

saliency of the current region. To avoid large value of the variance, the variance is normalized in the range of [0, 1]. The contrast ratio (RC) is a property measuring the ratio of the luminance of the brightest color (white) to that of the darkest color (black). A high contrast ratio represents a better color representation. We utilize the inverse contrast (CR) to measure the tendency of the smoothness of the relative intensity. The value of CR will be 0 for a constant area of a region, otherwise 1 for a high contrast area. The inverse contrast as follows,

$$CR = 1 - \frac{1}{1+\sigma_s^2} \qquad (14)$$

The dissimilarity between query regions and the regions associated with an image in the database can be measured based on the overall weighted distances [25]. These dissimilarities are used to establish a ranking list of the best candidate images. At this stage, We focus on adaptively refining the importance of the similarity from the search results to users' desired targets. This will cause irrelevant salient regions or non-salient regions on the object of interest to occur on the saliency map. These problems are tried to be solved by internal or external approaches to the saliency computation such as the normalization of feature maps (internal approach) and segmentation (external approach) to enhance salient object detection .

The Saliency map of each input channel enables us to obtain the saliency at decomposition level as,

$$F^c(x,y) = N(\sum_s f_s(x,y)) \qquad (15)$$

Where Fc is the saliency map of input channel c, and N (.) is the function that consists of Gaussian smoothing and normalization to the range [0, 1]. It focuses on global structures at its initial coarse level and gradually focuses on local details in the subsequent levels. Therefore, it can allow belief propagation to go across texture regions easily. The fuzzy vertex region algorithmic method gets stuck in texture-like regions in the background, while Gaussian smoothing can overcome these local textures and highlight the salient foreground objects successfully in these test images.

D. *Computational Complexity*

The computing complexity of MRF is not readily predictable. The computing time is proportional to the range for a belief to propagate from the source to the pixel is related to the number of pixels. Our experiment validated that the convergence time is nearly proportional to the number regions of pixels, namely O(N). It is noted that Fuzzy oriented saliency extraction algorithm needs to converge on multiple layers. However, the number of nodes has been reduced to $1/2^k$, and the search space at the top level of the pyramid is then drastically reduced to $(1/2^k)^2$ as well. While the pre-converged confidence map is inherited by the next level and hence it can be expected to help reduce the total convergence time. The computing time was measured for a MATLAB solution. We can see that the single-level took the longest time to converge (210 ms), for 5-level needs only 23 ms in total for all five levels. From this comparison successively not only helps overcome texture background, but also helps reduce the computing time.

IV.EXPREMENTS AND RESULTS

The aim of this experiment to test the ability to recognize the salient feature extraction ability of the proposed method. Fuzzy feature evaluation index helps to reconfigure the feature detection stage using the confidence values from fuzzy index value and a pre-defined minimum threshold.

E. *Exprimental Setup*

The similarity function GSI compares the saliency of the common features and the distinctive features to the saliency of the common features and the distinctive features. The values of a and b establish the relative importance of distinctive features in the similarity assessment. When $a \neq b$, a directional similarity measure is acquired and the similarity function is not symmetrical.

$$GSI(A,B;a,b) = \frac{f(A\cap B) + f(\bar{A}\cap\bar{B})}{f(A\cap B) + af(A-B) + bf(B-A) + f(\bar{A}\cap\bar{B})} \qquad (16)$$

Where f is a non-negative and increasing function, A, B denote the sets of true predicates on the measurements. To overcome the uncertainty of computational error during the reconfiguration process the Cauchy function can capture the feature vectors and reflect the retrieval accuracy as,

$$C(\vec{x}) = \frac{1}{1+\left(\frac{\|\vec{x}-\vec{v}\|}{d}\right)^a} \qquad (17)$$

$\vec{v}$ is the center point or location of the fuzzy set, *d* denotes the width and *a* determines the smoothness of the function, *d* and *a* portray the grade of fuzziness of corresponding fuzzy feature. For each region the fuzzy feature determined by

$$\tilde{F} = \frac{1}{2}\sum_{i=1}^{c-1}\sum_{k=i+1}^{C}\left\|\hat{f}_i - \hat{f}_k\right\|^a \qquad (18)$$

$\hat{f}_i$ is the average distance between the cluster centers and also describes the transition of the region boundaries.

F. *Dataset*

The evaluation and comparison of the proposed method against the previous baseline methods on two representative benchmark datasets, the MSRA 1000 and CIFAR100 are used. The MSRA-1000 dataset contains 1,000 images with the pixel-wise ground truth provided that each image in this dataset contains a salient object. CIFAR 100 is wildly used dataset consists of 100 classes of natural images. There are 50K training images and 10K testing images. Randomly cropped and flipped image patches of size 26 × 26 are used for training. The initial learning rate is 0.01 and it is decreased by a factor of 10 every 6K iterations. Fine-tuning is performed for 20K iterations with large mini-batches of size 256. The initial learning rate is 0.001 and is reduced by a factor of 10 once after 10K iterations.

Evaluate our saliency detection method by using two different datasets the MSRA1000 and CIFAR100 are used. Both datasets include images with complex background and low contrast objects, as well as manually labelled ground truth masks (GT) for salient object candidates. Several state-of-the-art saliency detection algorithms are chosen for comparison and are in the following referred to as saliency filter (SF), frequency tuned saliency(IG), Global Components (GC), and Region Contrast (RC) respectively. Before normalize all the saliency maps to the range of [0; 255] and obtain 256 binary results by varying the segmentation threshold. The precision and recall rates are firstly computed for each image at every threshold then averaged over the overall benchmarks.

*G. **Expremental results***

For each descriptor, the procedure is complete, the application arranges the images contained in the database according to their proximity to the query image, generating a ranking list. Overall, the system generates individual ranking lists. Then, using either the proposed method, the methods from the literature, these result lists are in order to generate the final ranking list. The measures used to evaluate the performance of the algorithms are False Negative Error (FNE), the False Positive Error (FPE), and the Weighted Miss Error (WME). Let us denote the set of ground truth points for model M as G (n; σ) and the set of interest points detected by the algorithm. Denoting the number of correctly detected points in G as NC and the number of points in G as NG, the false negative error rate at localization error tolerance r is defined as:

$$FNR(r) = 1 - \frac{N_c}{N_G} \qquad (19)$$

To calculate the false positive error rate, each correctly detected point g ϵ G corresponds to a unique a. All the points in A, excluding a correspondence in G, are considered as false positives. Then, the number of false positives, denoted as NF, is set to NF = NA - NC, where NA is the number of interest points detected by the algorithm. The false positive error rate at

Localization error tolerance r is then defined as:

$$FPR(r) = \frac{N_F}{N_A} \qquad (20)$$

To incorporate the prominence of an interest point into the evaluation, another miss error measure is defined, as the Weighted Miss Error (WME). The WME is defined as

$$WME(r) = 1 - \frac{1}{\sum_{i=1}^{n} N_G \delta_{pi} n_i}, \delta_{pi} = 1 \quad \forall g_i \qquad (21)$$

Where δ_{pi} is the degree of view point interest contributed by pi. Hence, this error intends to measure the ability of an algorithm to detect the semantically significant interest points. The WME has a range of 0 to 1 with the best matching quality defined by the value 0 and the worst by 1. Also the performance of the method using the Mean Average Precision (MAP) metric for evaluation:

$$percision(p) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \qquad (22)$$

$$recall(r) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ retrieved} \qquad (23)$$

The mean average precision is computed as follows

$$MAP = \frac{1}{Q}\sum_{q=Q}\frac{1}{N_R}\sum_{n=1}^{N_R} p_q(R_n) \qquad (24)$$

Where Rn is the recall after the nth relevant image retrieved and NR the total number of relevant documents for the query, Q is the set of queries q. additionally, to calculate significant the performance deviation between the methods by using Significance test. An observed effect, such as a difference between two means, or a correlation between two variables, could reasonably occur just by chance in selecting a random sample at significance levels 0.05, 0.01, and 0.001, against a baseline run.

V.PERFORMANCE EVALUATION

A. Performance on MSRA1000 dataset

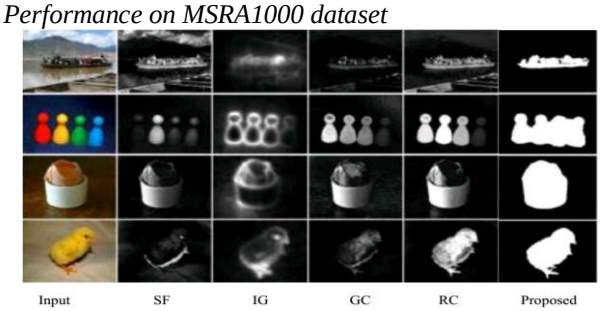


Figure 1. Visual Comparison of Saliency Map of Full Benchmark Dataset.

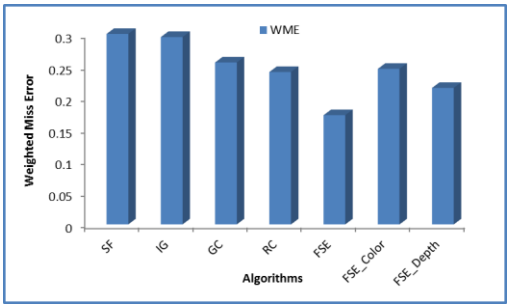
Figure 1 shows the computed saliency map compared the proposed global contrast based methods with 4 state-of-the-art saliency detection methods. They used our methods and the others to compute saliency maps for all the 1000 images in the database. Table 1 compares the average time taken by each method. For typical natural images, our FSE method needs O (N) computation time and is sufficiently efficient for real-time applications. In contrast, our RC variant is slower as it requires image segmentation, but produces superior quality saliency maps. The average time of each method is measured on a PC with Intel i7 3.3 GHz CPU and 8GB RAM. Performance of all the methods compared in this table is based on implementations in C++ and MATLAB.

TABLE 1 COMPARISON OF AVERAGE TIME TAKEN TO COMPUTE THE SALIENCY ON MSRA-1000 DATASET

TABLE I.

Algorithm	SF	IG	GC	RC	FSE
Time (S)	0.15	0.61	0.09	0.25	0.07
Code	C++	Matlab	C++	C++	Matlab

Figure 2 shows the performance of the five methods tested and ours, with respect to the localization error tolerance r. The lower the graphs, the better the results. With ideal interest point detection, the errors are expected to drop very quickly with respect to r. A rapid drop in WME means that the algorithm finds the interest points with a low localization error, while a rapid drop in WME indicates that the algorithm does not return excessive interest points.



In Table 2, present the significance test results at significance levels of 0.05, 0.01, and 0.001against the baseline. The proposed fuzzy method significantly improves the results, for all the three levels experimented with, and in all 4 evaluation measures. MAP value improved by 3.35% comparing to RC, 12.62% comparing to GC, 30.54% comparing to IG and 32.62% comparing to SF.

Table 2. Evaluation distribution of proposed method on MSRA-1000 Dataset

Algorithm	MAP	FPR	FNR	Precision	Recall
SF	0.3046	0.6101	0.3899	0.25	0.07
IG	0.3254	0.5124	0.4876	0.54	0.36
GC	0.5046	0.5376	0.1954	0.87	0.76
RC	0.5973	0.5175	0.4237	0.88	0.80
FSE	**0.6308**	**0.4300**	**0.1925**	**0.86**	**0.81**
FSE_Color	0.6147	0.4450	0.3962	0.78	0.69
FSE_Depth	0.5846	0.4550	0.3444	0.79	0.71

B. *Evaluation on CIFAR-100 dataset*

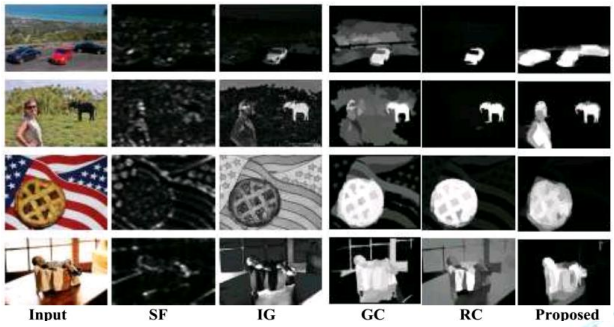


Figure 3. Visual Comparison of Saliency map between different algorithms on CIFAR-100 Dataset

In Figure 3 the dataset has 100 class natural images and the Pixel wise annotation of salient objects in this dataset was generated. This dataset is very challenging since many images contain multiple salient objects either with low contrast or overlapping with the image boundary.
To facilitate a fair comparison with other methods, randomly choose 10K images from the training set as the held-out set. Fine categories within the same coarse categories are visually more similar. It takes around 20 hours to train our deep prediction model for 15 image segmentation levels of regions using the CIFAR dataset. It only takes 8 seconds to detect salient objects in a testing image with 400x300 pixels on a PC with a 3.4GHz Intel processor using our MATLAB code.

Table 3 shows the impact of the use of shared layers on the weighted classification error, memory footprint, and the net execution time. The coarse category component and all fine category components use independent preceding layers initialized from a pretrained building block net. Most of these methods are more complex because they use demand high computational resources and large datasets. In contrast, our method is quite simple in both testing and training steps. The results obtained are still encouraging compared to other simple methods that obtain worse results. It is also interesting to analyze the first two values of the table. These results are obtained using non-convolutional networks and learning directly from the entire image. Like our baseline result, this poor performance confirms the advantage of learning from small regions, especially with complex data as it happens in this case.

COMPARISON OF TESTING ERRORS, MEMORY FOOTPRINT (MB) AND TESTING TIME (SECONDS) BETWEEN IMAGE BLOCKS ON CIFAR-100 DATASET

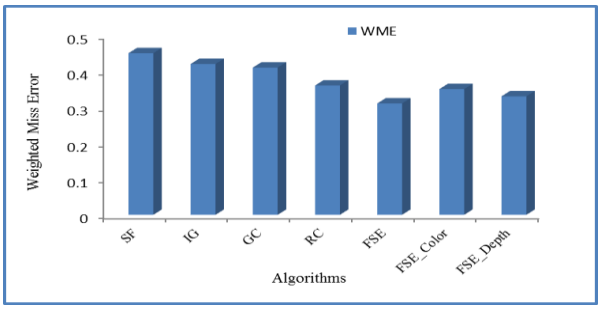


Figure 4: Comparison with State-of-art-in terms of Weighted Miss Error on CIFAR-100 Dataset

In Figure 4 shows that the localization error tolerance increases, our method detects more ground-truth interest points than the competing methods, having lower WME. Moreover, when multiple salient objects appear in an image, our model shows a dramatic advantage over all the methods mentioned above, which can locate individual region at pixel level within and on the contour of the objects exactly.For quantitative evaluation, to plot the precision-recall curves and F-measure in Figure 5 to compare our method with six state-of-the-art approaches. The precision and recall values clearly demonstrate that our model outperforms the other algorithms. Our model gets the highest precision value in almost the whole recall interval [0, 1], where the highest value can reach up to 0.93. The Proposed model can roughly outline the salient region of an image, the corresponding saliency values are not effectively enhanced. The quantitative comparison in Figure 5 shows the proposed algorithm performs well in both the precision-recall and the bar graph indicates the superior overall performance of our model.

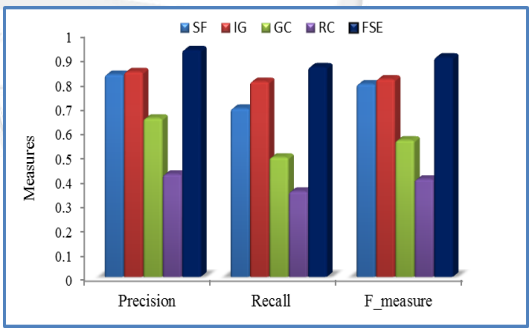


Figure 5. Quantitative Measurement Evaluation of Proposed Method FSE on CIFAR100 Dataset

VI.CONCLUSION

In this paper, a novel bottom-up computational model of visual attention is proposed to obtain the saliency map for images based on region coefficients. The proposed model derives feature maps of band-pass filtered regions from the input image with increasing frequency bandwidths. It can adapt the center-surround structure of the HVS (human visual system) since feature maps includes components from the edge to the texture based on the multi-level decomposition. Using these features, the local and global saliency maps are introduced to form the final saliency map. Saliency-on and saliency-off feature integration was proposed based on a fuzzy decision rule to demonstrate the reliability of saliency information on unsupervised segmentation model. Also, the

use of fuzzy logic to find automatic scale value is not used in the existing research.

VII.REFERENCES

[1] G. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, and Susstrunk S. "SLIC super pixels compared to state-of-the-art super pixel methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 11, pp. 2274–2281, 2012.

[2] Ali N, Bajwa K. B, Sablatnig R, and Mehmood Z. "Image retrieval by addition of spatial information based on histograms of triangular regions," Computers&Electrical Engineering, 2016.

[3] Ashraf R, Bashir K, Irtaza A, and Mahmood M. T. "Content based image retrieval using embedded neural networks with bandletized regions," Entropy, vol. 17, no. 6, pp. 3552–3580, 2015.

[4] Ashraf, Raja Shahid, Iain Meager, Mark Nikolka, Mindaugas Kirkus, Miquel Planells, Bob C. Schroeder, Sarah Holliday et al. "Chalcogenophene comonomer comparison in small band gap diketopyrrolopyrrole-based conjugated polymers for high-performing field-effect transistors and organic solar cells." Journal of the American Chemical Society 137, no. 3 pp: 1314-1321, 2015.

[5] Aurich, V., & Weule, J. Non-linear gaussian filters performing edge preserving diffusion. In Mustererkennung Springer Berlin Heidelberg. pp. 538-545, 1995.

[6] Bishop C.M., Pattern Recognition and Machine Learning, Springer, 2006.

[7] Borji A and Itti L. "State-of-the-art in visual attention modeling," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 185–207,Jan. 2013.

[8] Buades A, Coll B and Morel J.M. "A non-local algorithm for image denoising," Proc.IEEE CVPR, vol. 2, pp. 60-65, 2005.

[9] Cheng M.-M,. Zhang G.-X, Mitra N. J, Huang X, and Hu S.-M. "Global contrast based salient region detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11), pp. 409–416, June 2011.

[10] Chiu, C. Y., Lin, H. C., & Yang, S. N. (2003, May). A fuzzy logic CBIR system. FUZZ'03. The 12th IEEE International Conference on Fuzzy Systems, Vol. 2, pp. 1171-1176. 2003.

[11] Deming Kong, Liangliang Duan, Peiliang Wu, and Wenji Yang. "Salient Region Detection via Feature Combination and Discriminative Classifier" Mathematical Problems in Engineering, Volume 2015, pp.1-13, Nov 2015.

[12] Dickson, Monte Andre, Bingcheng Ni, and Shufeng Han. "Region of interest selection for varying distances between crop rows for a vision guidance system." U.S. Patent No. 6,490,539. 3 Dec. 2002.

[13] Duan L., Wu C, Miao J, Qing L, and Fu Y, "Visual saliency detection by spatially weighted dissimilarity," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 473–480, Jun. 2011.

[14] Eesa, A. S., Orman, Z., & Brifcani, A. M. A. A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. *Expert Systems with Applications*, vol 42 issue 5, pp.2670-2679, 2015.

[15] Erdem E and Erdem A. Visual saliency estimation by nonlinearly integrating features using region covariances. Journal of Vision, Vol 13, Issue 4 pp:1–20, 2013.

[16] Feng X, Liu T, Yang D, and Wang Y, "Saliency Based Objective Quality Assessment of Decoded Video affected by Packet Losses", Proc. of IEEE, ICIP, pp.2560- 2563, 2008.

[17] Fulkerson B, Vedaldi A, and Soatto S. "Class segmentation and object localization with super pixel neighborhoods". *In ICCV*, pages 670–677, 2009.

[18] Gal R and Cohen-Or D, "Salient geometric features for partial shape matching and similarity," ACM Trans. Graph., vol. 25, no. 1, pp.150-155, 2006.

[19] Huang X, Shen C, Boix X, and Zhao Q. "Salicon: Reducing the semantic gap in saliency prediction by adapting deep neural networks," in Proc. Int. Conf. Comput. Vision, pp. 262–270, 2015.

[20] Imamoglu N, Gomez-Tames J, Gonzalez J, Gu D, and Yu D., "PCNN Segmentation and Bottom-Up Saliency-On Feature Extraction for Thigh MRI based 3D Model Construction," Journal of Medical Imaging and Health Informatics, vol.4, pp.1-10, 2014.

[21] Jinliang Wu, Xiaoyong Shen, Wei Zhu, and Ligang Liu, "Mesh saliency with global rarity," Graphical Models, vol. 75, no. 5, pp. 255 – 264, 2013.

[22] Kim W, Jung C, and Kim C. "Spatiotemporal saliency detection and its applications in static and dynamic scenes," IEEE Transactions on Circuits System and Video Technology. vol. 21, no. 4, pp. 446–456, Apr. 2011.

[23] Konstantinidis K, Gasteratos A, Andreadis I. Image retrieval based on fuzzy color histogram processing. Optics Communications, vol.248 issue 4, 375-386, 2005

[24] Krig S. "Interest point detector and feature descriptor survey," in Computer Vision Metrics, Springer, Berlin,Germany, pp. 217–282, 2014.

[25] Leifman G, Shtrom E, and Tal A, "Surface regions of interest for viewpoint selection," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 414–421,2012.