© UNIVERSAL RESEARCH REPORTS | REFEREED | PEER REVIEWED
ISSN : 2348 - 5612 | Volume : 05 , Issue : 01 | January – March 2018

VIDEO SALIENCY DETECTION IN FREQUENCY DOMAIN USING MODIFIED HFT

***Yogita*[1], *Rainu Nandal*[2]**
[1]*M.Tech Research Scholar,MDU, Rohtak, Haryana*
[2] *Asst. Prof, M.Tech, MDU, Rohtak, Haryana*
riti_srm@yahoo.co.in, rainu_nandal@yahoo.com

Abstract- **As it is known that the ability of sensory system of human being to seek out salient object in image which is quite fast. However, machine modeling of this basic intelligent behavior still remains a challenge. Image as well as video saliency detection technique based on neural network is presented in this paper to distinct salient regions with their surroundings. Even though several existing saliency detection algorithms have been proposed whereas the obtained saliency maps have not given the satisfying results. In this paper the problem of visual saliency is tackled. First of all in this technique saliency detection is considered as a frequency domain analysis. Secondly, algorithm accomplished this by using the concept of non-saliency. Third, proposed algorithm tends to at an equivalent time consider the detection of salient regions of varied size. In this paper a replacement of bottom-up approach for detection of visual saliency, characterized by a scale-space analysis of the amplitude spectrum of natural pictures and videos. A modified Hypercomplex Fourier is proposed and experimental analysis demonstrates that model gives efficient results and achieved the accuracy of about 80%.**

Keywords: **Saliency Detection, Hypercomplex Algorithm, Gaussian Filter, Lucy Richardson Algorithm.**

ISSN : 2348-5612 © URR

I. INTRODUCTION

Of all our sense organs, our eyes get 80% of our day by day life data. This measure of visual information physically surpasses the limit of our cerebrum [1]. The system that beats this central issue and figures out what part of the approaching visual data is intriguing and must be handled first is called visual attention. The number of domains wherever prominence maps is used is increasing. This is often the case as an example in laptop vision, lighting tricks, human-computer interaction, neuromarketing, robotics, etc. a number of them area unit delineate hereafter such as Advertising analysis, Surveillance systems, Video Retargeting, etc [2]. Saliency discovery for pictures and recordings has turned out to be progressively prevalent because of its wide appropriateness as appeared in figure 1. In this exploration work, we propose a novel base up video saliency [3,4] extraction technique, which incorporates two sections: static saliency recognition and element saliency location. Firstly, we consider the static saliency identification as a grouping issue: a scene can be isolated to saliency locale and non-saliency area [5].

Figure 1: Saliency Detection

Saliency Detection is incredibly vital for image and video process application. Saliency algorithmic program is applied to the frames [7] of pictures to filter the background from the video frames.

The main drawback of getting objects from the scene and separating them from the background of image. The human brain will try this separation terribly simply and quick, however doing same on a machine is one in all the key challenges for engineers and scientists. Background separation from the photographs is that the one in all the task within the machine application makes it a vital a part of the system. The output of the Background separation is that the input to the higher-level method, as an example, the trailing of known objects from the image.

During the last decade, to get automatic, effective and correct saliency extraction ways, several researchers pay most attention on each static (image) saliency extraction. In saliency extraction, there are two models that is driven by: a bottom-up saliency and a top-down saliency. Bottom-up cues area unit chiefly supported characteristics of a visible scene (stimulus-driven), whereas top-down cues (goal-driven) area unit determined by psychological feature phenomena like data, expectations, reward, and current goals. In bottom-up ways, low-level options, like color, intensity and orientations, area unit computed to retrieve the fundamental components of distinction, that function the clues for light the foremost engaging regions in a picture[5-8]. For video saliency detection, saliency is taken into account as classification drawback that is solved by classification methodology. Once classification we have a tendency to extract saliency and non-saliency region of the video sequences. Then motion feature with the special feature to represent objects dynamic saliency in temporal domain [11,12].

II. RELATED WORKS

In [8] a region-based solution for saliency location material to better encode the picture features for settling object recognition task. To start with utilize the versatile mean move calculation to extricate superpixels from the information picture. At that point apply Gaussian mixture model (GMM) to group superpixels in view of their shading likeness. At long last, compute the saliency esteem for each group utilizing spatial minimization metric together with changed PageRank propagation.

In [9] model introduced a graph-based technique to figure visual salience. n the first place, indistinguishable element maps than inside the FSM demonstrate are extricated [9]. It winds up in 3 multiscale include maps: hues, power and introductions. At that point, a completely associated chart is made over all matrix areas of each element outline a weight is dispensed between every hub. This weight relies upon the reflection separate and along these lines the cost of the element outline hubs. At long last, every chart is dealt with as mathematician chains to influence AN actuation to delineate hubs that are to a great degree not at all like including hubs are apportioned high esteems. All initiation maps are joined into a definitive remarkable quality guide. Again here, exclusively privately differentiated choices are coordinated over the picture, the model is so mostly upheld native context.

In [10] a saliency recognition technique that recognizes the saliency of HDR pictures and HDR video outlines. The spatial and temporal signs are considered, prompting two saliency maps: the spatial saliency map and the temporal saliency map. In the first place, to acquire the spatial saliency delineate, the HVS model to break down element channels from a HDR info and after that take after the strategy of the bottom up approach. At that point to process the temporal saliency outline, optical flow based technique is utilized to find motion. Finally, a dynamic combination strategy is proposed to join both the spatial and temporal saliency maps.

In [11] author proposed an algorithm based on supervised learning technique to analyse and detect salient object or saliency map in image frames using conditional random fields (CRF). Most of the existing CRF techniques generates graphical pixel wise model using eight neighborhood grid-shaped graph. In this research paper superpixel level graph with two-ring with pseudo-background neighborhood technique is used for enhancing the performance of the system. The result analysis shows that the saliency maps generated have comparatively effective pure background regions.

In [13] creator proposed a novel base up video saliency extraction technique, which incorporates two sections: static saliency location and element saliency recognition. Firstly, we consider the static saliency discovery as an order issue: a scene can be partitioned to saliency locale and non-saliency area. To take care of such issue, we utilize a recurrence based inspecting technique and after that present a Canonical Correlation Analysis (CCA)- based order procedure utilizing numerous signs. Besides, we join movement include with the spatial element to speak to element saliency in worldly space.

Other than the saliency observation techniques above, we expect that saliency recognition can be taken as a sifting procedure, which is a performed in the recurrence space to sift through the normal vitality flag and hold different features of the incorporated flag vitality contrast bigger spectral separating process. Along these lines, we propose a novel spatio-fleeting saliency observation technique in light of hypercomplex spatiotemporal spectral contrast (HSC). Also demonstrating the heartiness and viability of proposed techniques, it has been stretched out the application to moving article i.e. in video.

III. PROPOSED METHODOLOGY

The center of this research work is on one of our visual capabilities- discovering objects of interest for pictures and in addition recordings. There are two principle features of this issue - finding an unknown salient object and finding a specific known object. In this paper, we name the principal issue of finding an obscure question of intrigue saliency identification and the second one of recognizing a known protest of intrigue assignment particular protest location or just protest discovery. For the previous, we create saliency discovery calculations.

A. Hypercomplex Fourier Transform

The contribution to the customary Discrete Fourier Transform is a genuine grid. Every picture pixel is a component of the information grid and is a genuine number. Notwithstanding, in the event that we consolidate more than one component into a hypercomplex lattice, every component is a vector and this hypercomplex grid is a vector field. Along these lines, the customary Fourier Transform gets to be unsatisfactory for computational purposes. The adjusted Hypercomplex Fourier Transform is proposed, in which the hypercomplex info was determined to be a quaternion5. Given a hypercomplex lattice:

f(n,m) = a + bi + cj + dk,

the discrete version of the HFT of is given by:

$$\mathrm{FH}[u,v] = 1/\sqrt{\mathrm{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu 2\pi\left(\left(\frac{mu}{M}\right)+\frac{nu}{N}\right)} \mathrm{f}(n,m)$$

where μ is a unit pure quaternion and μ2 =-1. Note that FH[u, v] is also a hypercomplex matrix. The inverse Hypercomplex Fourier Transform is given as:

$$\mathrm{f}(n,m) = 1/\sqrt{\mathrm{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu 2\pi\left(\left(\frac{mu}{M}\right)+\frac{nu}{N}\right)} \mathrm{FH}[u,v]$$

Hypercomplex Representation of Multiple Feature Maps

The Hypercomplex representation can be utilized to join different components as shading, power and movement as the elements. We characterize the information hypercomplex [14] network as takes after:

f(n,m) = wafa + wbfbi + wcfcj + wdfdk

where wa-wd are weights and fa-fd are the feature maps (matrices) and:

fb = (r + g + b)/3,

fc = R − G,

fd = B − Y,

where, r, g, b are the red, green and blue channels of an input color image and

R = r−(g+b)/2,

G = g−(r+b)/2,

B = b − (r + g)/2, Y = (r + g)/2 − |r − g|/2 − b.

These three element maps contain the adversary shading space representation of the info picture our approach has additionally been tentatively affirmed utilizing recordings by characterizing a movement include M and setting f1 = M. In this exploration work, we consider the static picture case and movement instance of recordings by utilizing just power and shading data. We select the weights so that w1 = 0.5, w2 = 0.5, w3 = w4 = 0.3.

B. Algorithm Steps

Input:
Video
Output:
Salient object detection in video frames
1: Frame extraction out of input video.
2: Afterward extraction of image frames from video we take individual frames and enhance the given image using Lucy-Richardson Algorithm is used for noise removal and de-blurring the blurred image. Lucy Richardson (LR) algorithm is an iterative nonlinear restoration technique.
3. Compute the feature vector or maps.
4: Form the hypercomplex matrix f(n,m) by combining these feature maps
5: Perform the Hypercomplex Fourier Transform on f(n,m) and compute the amplitude spectrum A, phase spectrum, P and eigen axis spectrum.
6: Smooth the amplitude spectrum with Gaussian kernels thereby obtaining a spectrum scale space.
7: Obtain a saliency map Sk for each frames and thereby producing a sequence of saliency maps.
8: Find the best saliency map S using neural network from the feature map set.

ISSN : 2348 - 5612 | Volume : 05 , Issue : 01 | January – March 2018

9: Final Salient object S.

IV. RESULTS

We first evaluate the performance of our static saliency extraction. We test our system using 100 images of different types as well as videos which has been widely used in evaluating image saliency detection algorithms. The testing image size is 400*300 pixels. We implement our method using following to show the performance of our static saliency extraction using MATLAB. Figure 2-6 represents the whole process for finding salient region out of video frames. Table 1 represents the total time elapsed in finding the salient object out of video frames.

Figure 2: Image frames out of video

Figure 3: Enhanced Image

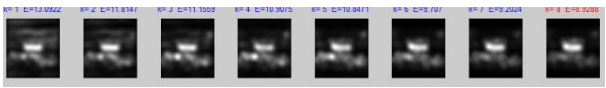

Figure 4: Saliency map set

Figure 5: Best Salient object detected

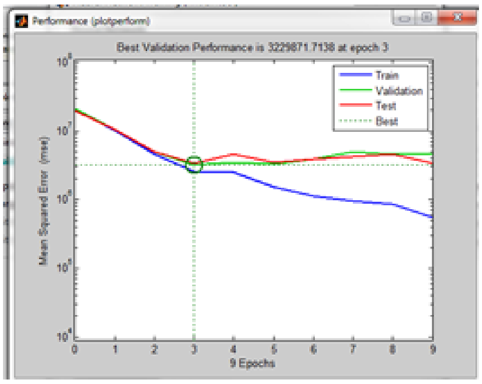

Figure 6: Mean Squared Error

Table I: Tabular representation of result data

Fig. No.	Time in calculation
1	1.043885 seconds
2	0.965437 seconds
3	0.920276 seconds
4	1.005192 seconds
5	0.993353 seconds

A. Performance Evaluation

To evaluate the performance of the proposed system following parameters such as Precision, Recall and Fmeasure are used.

Precision = TP/ (TP+FP)

Recall = TP/(TP+FN)

Accuracy= (TP+TN)/(TP+TN+FP+FN)

Fmeasure = 2* (Precision* Recall) / (Precision + Recall)

Where, True Positive (TP) = Correctly detected object in image.

© UNIVERSAL RESEARCH REPORTS | REFEREED | PEER REVIEWED
ISSN : 2348 - 5612 | Volume : 05 , Issue : 01 | January – March 2018

True Positive (TN) = No object region correctly detected in image.
False Positive (FP) = Object incorrectly identified in images.
False Negatives (FN) = Object that are failed to be detected in image.
The results of detected phase and recognition phase of proposed methodology are evaluated on 10 different video frames and the result analysis of these frames are illustrated in Table II to VI. To evaluate the performance of the saliency detection on the parameters such as Precision, Recall, Accuracy, Fmeasure and Time are compared are illustrated in figure 7 to 12.

Table II: Recall Analysis

S. NO.	VIDEO NO.	RECALL
1	VIDEO-1	1
2	VIDEO-2	1
3	VIDEO-3	0.8922
4	VIDEO-4	1
5	VIDEO-5	0.9871
6	VIDEO-6	1
7	VIDEO-7	0.9755
8	VIDEO-8	0.7392
9	VIDEO-9	0.6796
10	VIDEO-10	1

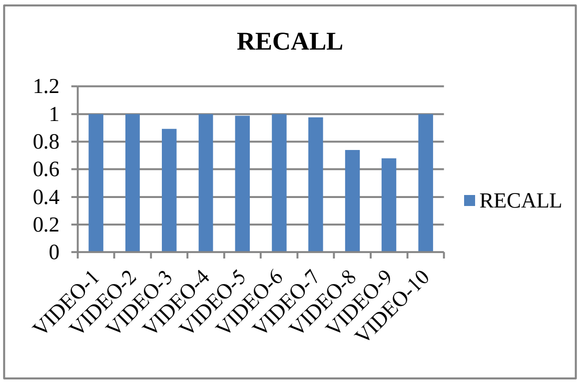

Figure 7: Recall Analysis

Table III: Precision Analysis

S. NO.	VIDEO NO.	PRECISION
1	VIDEO-1	0.7802
2	VIDEO-2	0.7097
3	VIDEO-3	0.871
4	VIDEO-4	0.9703
5	VIDEO-5	0.9029
6	VIDEO-6	0.9609
7	VIDEO-7	0.9447
8	VIDEO-8	0.8613
9	VIDEO-9	0.6637
10	VIDEO-10	0.9192

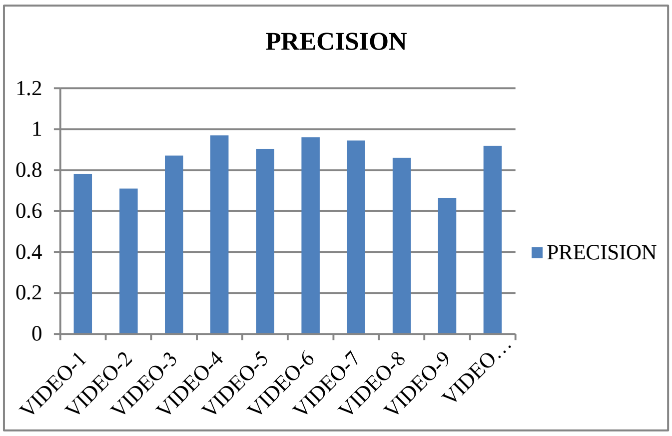

Figure 8: Precision Analysis

Table IV: Accuracy Analysis

S. NO.	VIDEO NO.	ACCURACY
1	VIDEO-1	0.7802
2	VIDEO-2	0.7097
3	VIDEO-3	0.7881
4	VIDEO-4	0.9703
5	VIDEO-5	0.8924
6	VIDEO-6	0.9609
7	VIDEO-7	0.9229
8	VIDEO-8	0.6606
9	VIDEO-9	0.5055
10	VIDEO-10	0.9192

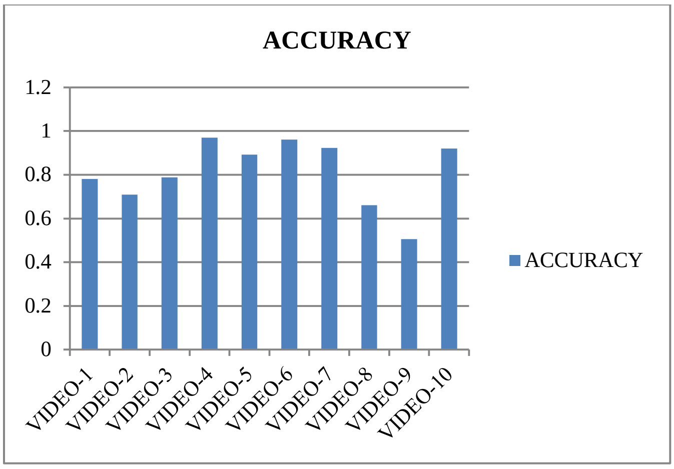

Figure 9: Accuracy Analysis

Table V: F_Measure Analysis

S. NO.	VIDEO NO.	F_MEASURE
1	VIDEO-1	0.8765
2	VIDEO-2	0.8302
3	VIDEO-3	0.8815
4	VIDEO-4	0.9849
5	VIDEO-5	0.9431
6	VIDEO-6	0.9801
7	VIDEO-7	0.9599
8	VIDEO-8	0.7956
9	VIDEO-9	0.6715
10	VIDEO-10	0.9579

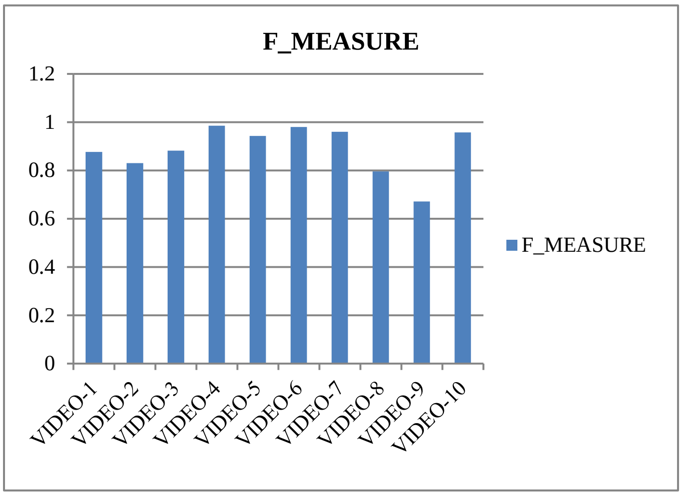

Figure 10: F_Measure Analysis

Table VI: Time Analysis

S. NO.	VIDEO NO.	TIME
1	VIDEO-1	1.1264
2	VIDEO-2	1.156
3	VIDEO-3	1.38
4	VIDEO-4	1.13
5	VIDEO-5	2.312
6	VIDEO-6	1.11
7	VIDEO-7	1.215
8	VIDEO-8	1.322
9	VIDEO-9	1.122
10	VIDEO-10	1.36

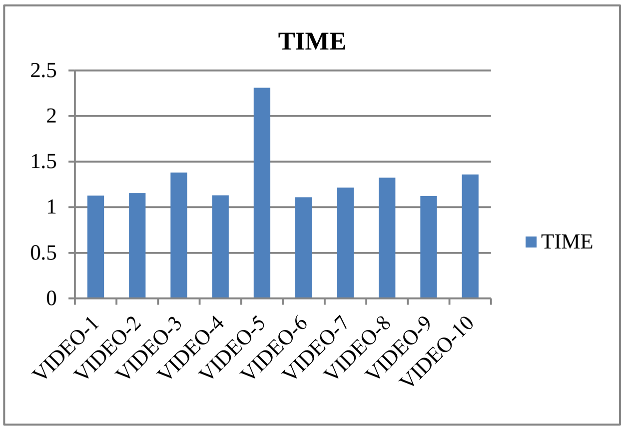

Figure 11: Time Analysis (in sec)

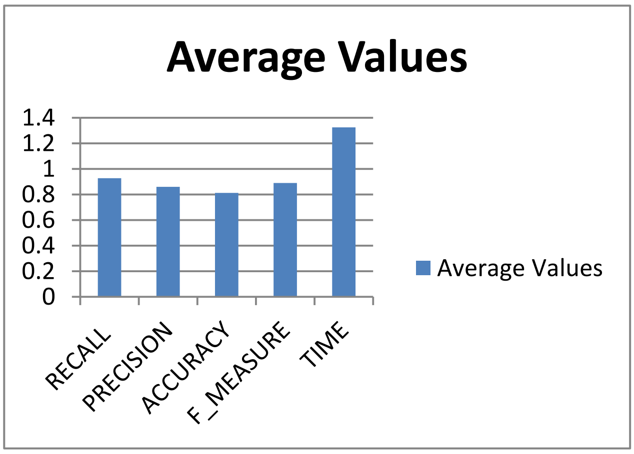

Figure 12: Chart for Average Performance Measure Values

V. CONCLUSION

This research work suggests a unique saliency detection frame for videos and image frames, based on analyzing the spectrum scale-space. In this paper a image or video saliency detection technique is presented that demonstrate the image convolution with a low-pass Gaussian kernel of an appropriate scale. The proposed methodology is capable to find minor and bulky salient objects in video frames. In direction to fuse multi-dimensional feature maps of video frames, a modified neuro-based Hypercomplex Fourier transform is demonstrated for spectrum scale-space analysis. To demonstrate the proposed approach efficiency, saliency computations is performed on different videos as well as natural images, and so compared the results with existing algorithms. So, as to form a fair comparison the proposed work using the MSE.

REFERENCES

[1] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient region detection and segmentation," In International Conf. On Computer Vision Systems, 2008.
[2] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," In IEEE CVPR, p. 1597C1604, 2009.
[3] A. Borji, "Boosting bottom-up and top-down visual features for saliency estimation," In IEEE CVPR, 2012.
[4] A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: A benchmark," In Proc. European Conference on Computer Vision (ECCV), 2012.
[5] Ming-Ming Cheng, Jonathan Warrell, Wen-Yan Lin, Shuai Zheng, and Vibhav Vineet, "Efficient salient region detection with soft image abstraction," In ICCV, 2013.
[6] Bowen Jiang, Lihe Zhang, Huchuan Lu, and Minghsuan Yang, "Saliency detection via absorbing markov chain," In ICCV, 2013.
[7] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang, "Saliency detection via graph-based manifold ranking," In IEEE CVPR, 2013.
[8] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," In ICPR, 2004.
[9] C. K. J. Harel and P. Perona, "Graph-based visual saliency", Proceedings of Neural Information Processing Systems (NIPS), 2006.
[10] Yuanyuan Dong, Mahsa T. Pourazad, and Panos Nasiopoulos, "Human Visual System-based Saliency Detection For HDR Content, " IEEE Transactions On Multimedia, Vol. 18, No. 4, April 2016.
[11] Wenliang Qiu Xinbo Gao Bing Han, "A superpixel-based CRF saliency detection approach", Elsevier, 2017.
[12] B. Schauerte and R. Stiefelhagen, "Predicting human gaze using quaternion dct image signature saliency and face detection", In Proceedings of the 12th IEEE Workshop on the Applications of Computer Vision (WACV) / IEEE Winter Vision Meetings, Breckenridge, pp. pp 9–11, January 2012.
[13] Kang Xue et al., "A Video Saliency Detection Method Based on Spatial and Motion Information", ICIP, IEEE 2015.
[14] Jian Li, Martin D. Levine, Xiangjing, Xin Xu and Hangen He, "Visual Saliency Based on Scale-Space Analysis in the Frequency Domain", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007.