Webpage: www.ijaret.org

Volume 3, Issue VII, July 2015
ISSN 2320-6802

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS…..

Enhanced Saliency Detection for 3D Multimedia Images

Y. Sri Chandana Gayatri[1], B. Doss[2]

[1]PG student, Dept of ECE, JNTUA College of Engineering,
Anantapur, Andhra Pradesh, India.
srichandana.gayathri@gmail.com
[2]Lecturer, Dept of ECE, JNTUA College of Engineering,
Anantapur, Andhra Pradesh, India.

Abstract: *Saliency detection is believed to be an important precursor for a vast number of multimedia processing applications. During the past decades, even though several saliency detection models have been introduced for many multimedia processing applications, there is a great need for an efficient saliency detection model for a vast number of new applications of stereoscopic display emerging at present. Besides 2D saliency detection methods, depth feature is also considered to detect saliency in stereoscopic images. This paper presents an enhanced saliency detection framework for stereoscopic images, depending on the feature contrast of luminance, color, texture and depth, which in turn are extracted from discrete cosine transform coefficients for the purpose of feature contrast calculation. Here, in order to consider the calculation of local and global contrast, a Gaussian model of spatial distance between image patches is adopted. In addition, this paper newly incorporates Hidden Markov model to improve the efficiency by mainly detecting the most salient regions in stereoscopic images. Later to obtain the resultant saliency map, the feature maps are combined in a unique fusion method and the two vital characteristics of human visual system that are center bias factor and human visual acuity, are adopted to enhance the resultant saliency map. This enhanced method based on the Hidden Markov model, exhibits better performance compared to the other methods implemented in existing studies.*

Keywords: *Stereoscopic images, saliency detection, multimedia processing applications, human visual system, hidden markov model.*

1. INTRODUCTION

One of the important characteristics of the human visual system for visual information processing is Visual attention. To lessen the complexness of scene analysis, Visual attention would selectively process the significant part out of vast amount of visual information, by filtrating others. In natural images, this significant visual information is also called as salient regions or Regions of interest (ROIs). In visual attention mechanism, there are two different approaches: one is top-down approach, which is task dependent cognitive processing affected by the performed tasks, feature distribution targets, etc., while the other is bottom –up approach. This is data driven and task independent perception process for automatic salient region detection for natural scenes. At present, there are plenty of bottom-up models to detect saliency in 2D images and videos.

Several efforts were made in the past by many studies, to find out computational methods of visual attention for diverse multimedia processing applications (e.g., Visual retargeting [3]). These applications specifically processed the extracted salient regions since they pull most of our attention towards them when compared to other locations. The advancement in stereoscopic display is the main cause for diverse burgeoning 3D multimedia applications today. Chamaret *et al.* employs ROIs in the study [9] for the purpose of 3D rendering. The rising demand of applications based on visual attention amplifies need of computational models of visual attention for 3D multimedia data.

Only some studies for 3D saliency detection exist today when compared to vast number of traditional 2D saliency detection models. To detect saliency in 3D images, we consider binocular depth cues and fuse them together with others (like monocular disparity) in a way that is pertinent to different viewing space conditions. These binocular depth cues are considered to achieve depth perception which has a great impact on the human viewing behavior.

In an eagle's view, it is obvious that the two factors that play an important role, to design 3D saliency detection models are: estimation of saliency from depth cues and combining this saliency with those obtained from other 2D low-level features.

This paper introduces an enhanced saliency detection method particularly for 3D images by considering color, texture, luminance and depth feature extracted from DCT (discrete cosine transform) coefficients of image patches. This enhanced model mainly employs DCT coefficients for extracting features of image patches, since it is well known for its energy compaction job, representing maximum amount of signal information, only on some low frequency components. This energy compactness property of DCT is the reason for employing it extensively in many signal processing applications in the past. One of the previous studies [3] also reflects its importance in detecting saliency.

The main tasks performed here, in detecting saliency are abridged below:

At the outset, the considered stereoscopic image and its depth map are brought in to small image patches. From each image patch of original image, we extract the corresponding DCT coefficients. From these DCT coefficients, we extract luminance, color and texture features, while the depth map's DCT coefficients help in extracting depth feature.

To consider local and global contrast, center-surround feature difference is used for feature contrast calculation including a Gaussian method of spatial distances between the patches of image, to weight it. Now the saliency map can be obtained by

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

implementing a unique fusion method for combining feature maps. The previous methods to detect the saliency in 3D images, either take up depth information to weight the customary 2D saliency map, or, the depth map and the existing 2D saliency map are simply combined to extract the saliency map in case of 3D images. Unlike those traditional methods, in this model, along with the luminance, color and texture features, depth feature is also considered simultaneously for saliency calculation in an entire framework and a peculiar fusion method is implemented to get the saliency map. Further this saliency map is enhanced by incorporating center bias factor and human visual sensitivity [13], which are the crucial characteristics of Human Visual System. In addition, a robust Hidden Markov model is incorporated additionally to improve the efficiency of the model by identifying the salient regions possessing higher saliency than other regions, in the final saliency map. Experiments conducted on few databases for stereoscopic images, reflect the better performance of this enhanced model compared to other existing ones.
The rest of the paper is structured in the following manner. The work related to the literature is introduced in the section II. Section III consists of detailed description of the enhanced saliency detection model. Section IV comprises results from experiments on a few databases for stereoscopic images. Finally section V gives the conclusion.

2. LITERATURE REVIEW

We are well-known with the fact that several computational methods of visual attention for a plethora of 2D multimedia processing applications were present in the past. Itti *et al.* considered neuronal architecture of order primates as the important basis and proposed a conservative visual system [2] that has feature contrast calculation from orientation, color and intensity as a plinth of its saliency detection process. A graph-based concept by Harel *et al.* is an extension to Itti's model. This theory measures saliency from feature difference more accurately than that of Itti's. Mainly considering the characteristics of the human visual system like contrast sensitivity operations, visual masking, perceptual decomposition, and also center-surround phenomenon, Le Meur *et al.* introduced a computational method of visual attention [6]. Recently patch-based contrast, proposes few models to detect saliency yielding promising results in extracting salient regions [3]-[5].
Besides all these models detecting saliency in 2D pedestal, many works processed 3D multimedia data to detect saliency. In study [7], a stereo attention structure introduced by Bruce *at al.*, is the extension of an old attention framework to the binocular environment. But it has no computational method. Taking multiple perceptual stimuli in to account, a visual attention method for 3Dvideo, based on stereoscopic vision was introduced by Zhang *et al.* [8]. Some studies use depth maps in order to weight the maps obtained through 2D saliency detection [8], while, some other studies incorporate depth map in to the conventional 2D saliency detection methods [10], to obtain saliency maps for 3D multimedia content. The color feature and the depth feature was employed in designing saliency detection method for 3D content, by Ciptadi *et al.* for image segmentation task in the study [12]. Lately, Wang *et al.* made an extension to the ancient saliency detection for 2D imagery and presented a computational 3D visual attention structure. Finally these works provide the key to accurately detect the saliency for 3D imagery by considering depth factor during saliency detection process.

3. ENHANCED MODEL

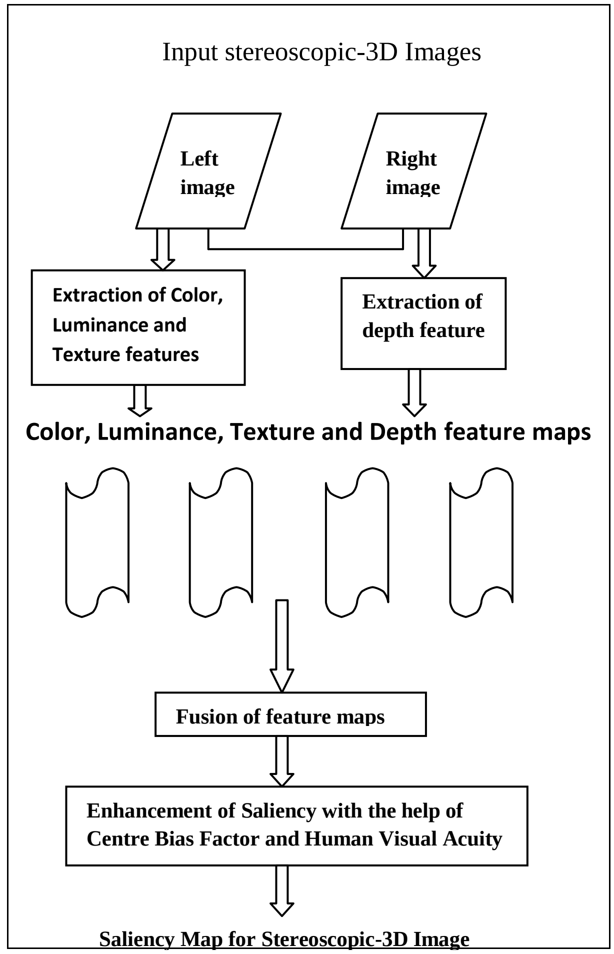

Fig.1: The framework of the enhanced saliency detection model

The underlying structure of the enhanced model introduced in this paper is represented by Fig.1. The initial stages of the model involves, extraction of depth feature along with few 2D features chosen as, luminance, color, and texture, from a stereoscopic image given as input. Corresponding feature maps are calculated based on their feature contrast. A saliency map is obtained by the combination of these feature maps using a special fusion process. As additional factors, Center bias and human visual acuteness (human visual system's important characteristics) are included in the model for enhancement of obtained saliency. Newly a robust Hidden Markov model applied to this saliency map can exhibit better

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

performance in detecting the regions possessing high saliency, when compared to the earlier models.

3.1 Extraction of Features

Based on the previous studies, we adopt the DCT coefficients for representing the energy of each path after dividing the considered input image in to small patches of size 8×8 whose size is equal to the block size of DCT, in compressed images of JPEG to give better performance. Conversion of RGB (image input) in to YCbCr color space, is performed since it possess perceptual property. Y in the YCbCr term stands for information regarding luminance, whereas Cb and Cr components are color opponent. The overall average energy for all pixels of the image patch is given by DC coefficients of DCT, whereas its AC coefficients exhibit the image patch's frequency properties in detail. From the above fact known, Y component's DC coefficients are used in representing the luminance feature i.e. *L =YDC* where *YDC* implies Y component's DC coefficient and Cb and Cr component's DC coefficients are used in representing the color features i.e.$C1$ =*CbDC* (DC coefficients of Cb) and $C2$ =*CrDC* (DC coefficients of Cr). The texture feature of the each patch is obtained from AC coefficients of Y component only, and not those of Cb and Cr, since they do not possess much information regarding texture. From the study [15], we consider only the first 9 AC coefficients of low frequency in representing each image patch's texture feature as, T = {*YAC*1,*YAC*2, . . ., *YAC*9}.

A disparity map exhibiting each pixel's parallax, between the images of left-view and right-view, represents depth information, and is measured mostly in pixel units in case of display systems. This study computes the depth information mainly considering the disparity as:

$$M = V/(1 + \frac{d.H}{P.W}) \quad (1)$$

Where V indicates observer's viewing distance; P denotes the disparity present between image pixels; H and W symbolize display screen's horizontal resolution and width (cm), respectively. All the values of these parameters are set depending on experimental works in the study [11]. The DC coefficients of the depth map's (M's) image patches are considered to extract the depth feature as D=MDC, where MDC stands for DC coefficient of the depth map's image patch.

As in above description, the extracted features (L, C1, C2, T, D) for a given stereoscopic image, play an important role in the calculation of corresponding feature maps and is introduced in the following subsection.

3.2. Feature Contrast Calculation

The saliency of the regions present in visual scenes can be given by feature contrast of each region from its surroundings. So, to obtain saliency, feature contrast must be calculated between each image patch all the remaining patches present in the image. The feature contrast is weighted by a Gaussian model of spacial distance between each patch of image and rest of the patches. F_i^k indicates the saliency value as:

$$F_i^k = \sum_{j \neq i} \frac{1}{\sigma\sqrt{2\pi}} e_{ij}^{l^2/(2\sigma^2)} . U_{ij}^k \quad (2)$$

where i and j represents image patches; k denotes feature and $k \in \{L, C1, C2, T, D\}$; the feature contrast from feature k, between i and j image patches is represented by U_{ij}^k; l_{ij} symbolizes spatial distance between the patches (i and j). with respect to the experimental work in previous study [3], the parameter (σ) of Gaussian model, is assigned the value 5, where it gives a measure of local and global contrast essential for estimating saliency. Besides feature contrast, the center-surround difference, weighted by spatial distances (within the Gaussian model) between each image patch and rest of the patches in the image, is needed for saliency calculation. The nearer image patches yield high contribution to the patch i's saliency value compared to the image patches far away from it. This phenomenon helps in considering local and global contrast from the obtained features in this current model.

As said earlier, DC coefficient for each patch represents the color, luminance and depth features. The difference between DC coefficients of image patches i and j results in the feature contrast from color, luminance and depth features between i and j and is mathematically represented as follows:

$$U_{ij}^m = \frac{|B_i^m - B_j^m|}{|B_i^m + B_j^m|} \quad (3)$$

where, feature is represented by B^m and $B^m \in \{L, C1, C2, T, D\}$; and feature contrast is normalized by introducing $B_i^m + B_j^m$ in the denominator. The L2 norm is used to calculate feature contrast $U_{ij}^{'}$ between two patches I and j from texture feature since 9 AC coefficients of low frequency are used to represent it. $U_{ij}^{'}$ is calculated as shown below.

$$U_{ij}^{'} = \frac{\sqrt{\sum_t (B_i^{'t} - B_j^{'t})^2}}{\sum_t (B_i^{'t} + B_j^{'t})} \quad (4)$$

Where AC coefficients are represented by t and t ∈{1, 2, ... , 9}; texture feature by $B^{'}$; and feature contrast is normalized by the denominator in Eq.(4).

3.3. Further Enhancement in detecting Saliency through Hidden Markov Model (HMM)

A more enhanced saliency detection technique is introduced here using Hidden Markov model which has its wide influence in many applications of image processing. A Markov process is a system which is modeled in a statistical Markov model having unobservable states, and can be called as a Hidden Markov model. The output based on the state is observable but the state cannot be observed directly. A probability distribution function exists for every state over the probable outputs.

Here we implement this technique in order to detect the most probable salient locations among all the possible salient regions detected. To fulfill this, the three tasks generally performed using a Hidden Markov method, are applied to this saliency detection model to increase the efficiency of the model. Generally in this model, the likeliness of the most salient locations associated with the observed salient regions can be computed as follows:

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

$$F_k^* = \text{argmax}_s \text{P}(\text{m}|\text{r}, \text{F}) \qquad (5)$$

where F^* represents the value of each of the most salient location of an image; the phrase of the Eq. (5), i.e. $(\text{m}|\text{r}, \text{F})$ gives the most probable salient regions among salient regions associated with all the patches present in the image.

3.4. Estimation of Saliency from Fusion of Feature Maps

The feature maps of color, luminance, texture and depth obtain above are fused in order to estimate the saliency map. A visual scene's saliency can be obtained through, interaction and simultaneous contribution of several visual features to it. Unlike the fusion process in the many previous works (e.g.[11]) that linearly combine feature maps in a simple way weighting the combination with constant values for all the feature maps, this paper uses adaptive weighting [1] for the feature map fusion to overcome the defects of adhoc weighting. The adaptive weighting assigns higher weighting for feature maps consisting of small and compact salient locations and the maps with widely spread salient areas are less weighted, since a saliency map is said to be efficient when its feature maps detect even smaller and compact locations of the image. A feature map exhibiting lesser compactness has larger spatial variance. The compactness β_k for a feature map F_k can be calculated by employing normalized values of spatial variance ϑ_k and is:

$$\beta_k = 1/(e^{\vartheta_k}) \qquad (6)$$

In the Eq. (6), k stands for feature channel and is, $k \in \{\text{L}, C_1, C_2, T, D\}$; ϑ_k represents the spatial variance and can be computed as follows:

$$\vartheta_k = \frac{\sum_{(i,j)} \sqrt{(i-E_{i,k})^2 + (j-E_{j,k})^2} . F_k^*(i,j)}{\sum_{(i,j)} F_k^*(i,j)} \qquad (7)$$

where feature map's spatial location is represented by (i, j); the spatial locations-average is weighted by feature response and is computed as:

$$E_{i,k} = \frac{\sum_{(i,j)} i . F_k^*(i,j)}{\sum_{(i,j)} F_k^*(i,j)} \qquad (8)$$

$$E_{j,k} = \frac{\sum_{(i,j)} j . F_k^*(i,j)}{\sum_{(i,j)} F_k^*(i,j)} \qquad (9)$$

Now the feature maps are fused depending on the property of compactness obtained in Eq. (6), in the following way.

$$S_f = \sum_k \beta_k . F_k^* + \sum_{p \neq q} \beta_p . \beta_q . F_p . F_q \qquad (10)$$

Where each feature map's compactness property weights it, in the first term representing linear combination of those maps; while, the enhancement of common salient regions is performed in the next term considering any two dissimilar feature maps.

The Fig.2 (h) reflects the adaptive weighting phenomenon and assigns higher weighting to the depth map during future map fusion resulting in detecting higher salient regions more accurately.

3.5. Enhancing the saliency map

To enhance the obtained saliency map, we consider centre bias phenomenon [1] that is independent of, distribution of the features in images.

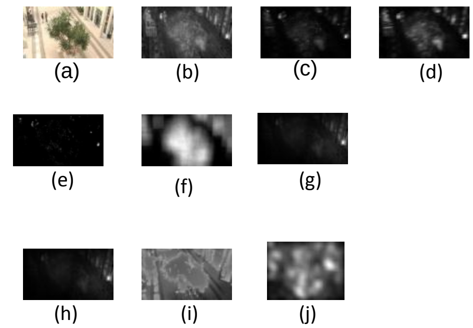

Fig.2: Examples for the feature maps of considered features and saliency maps: (a) original image; (b) color feature map from Cb component; (c) color feature map from Cr component; (d) luminance feature map; (e) texture feature map; (f) depth feature map; (g) saliency map through linearly combining feature maps with the constant weighting; (h) saliency map through the adaptive weighting process (the weights of Cb, Cr, L and D features maps are 0.45, 0.51, 0.49, 0.62, and 0.81, respectively); (i) enhanced saliency map using HMM; (j) ground truth map.

According to study [1], a special Gaussian function is used to obtain centre bias factor CBM denoted by S_c. Taking centre bias in to consideration, we calculate the saliency as:

$$S = \gamma_1 S_f + \gamma_2 S_c \qquad (11)$$

where the parameters γ_1 and γ_2 are employed in weighting saliency map S_f and centre bias map S_c and are assigned with 0.7 and 0.3 values respectively, Since we give higher importance to the saliency map than to the centre bias map S_c. The salient regions gain focus of human eyes during observation of natural scenes. The human visual acuity gets decreased in case of neighboring regions that are far away from these salient locations. From the study [13], saliency map is weighted by human visual sensitivity. The calculation of contrast sensitivity $C_s(f,e)$ is as follows:

$$C_s(f,e) = \frac{1}{C_o \exp\left(\alpha f (e+e_2)/e_2\right)} \qquad (12)$$

Where f represents spatial frequency in (cycles/degree); e represents retinal eccentricity in (degree); $C_o$represents minimum contrast threshold; α stands for spatial frequency decay constant; e_2 represents half resolution eccentricity. From study [13], the values that are best fit to the parameters α, e_2, C_o are 0.106, 2.3, and 1/64 respectively. In Eq. (12), the retina eccentricity (e) for a pixel location (i, j), accounting spatial distance between the pixel (i, j) and its closest salient pixel (i_o, j_0), is given by:

$$e = tan^{-1}(d'/\vartheta) \qquad (13)$$

Where, ϑ represents viewing distance; d' represents distance between (i_o, j_0) and (i, j) in the spatial domain. The normalized visual sensitivity $C_s(f,e)$ enhances the final saliency map S' that is computed as:

$$S' = S * C_s(f,e) \qquad (14)$$

ISSN 2320-6802

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS…..

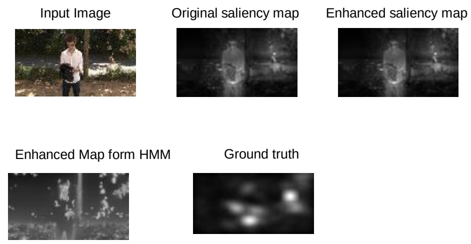

Fig. 3: Comparison examples between original and enhanced map using human visual acuity and center bias factor.

The inclusion of center bias factor increases the value of salient regions existing at the center of the image and human visual acuity decreases the saliency value of the regions that are non salient in natural scenes. The Fig. 3 compares original saliency map and the saliency map after enhancement through center bias factor and using human visual acuity, and also provides the enhanced map obtained from the inclusion of HMM in this model by considering a sample image.

4. EVALUATION FROM EXPERIMENTS

The section provides the performance of this enhanced model by conducting experiments. Later we compare different feature maps in subsection IV-2. The proposed method's is compared with other methods existing in previous studies and presents their performance evaluation. Some of the images considered here are taken from *Middlebury 2005/2006 dataset* [14]. According to the study [1][11], the quantitative performance of the introduced model in this paper is evaluated by considering the PLCC (Pearson Liner Correlation Coefficient), and KLD (Kullback-Leibler Divergence) along with AUC (Area Under the Receiver Operating Characteristics Curve). If the values of PLCC and AUC are larger, it indicates the efficiency of the model in predicting more accurate salient regions for 3D scenes. On the contrary, smaller the KLD values are, much better is the performance of the model.

TABLE-I

COMPARISON OUTCOMES OF PLC, KLD AND AUC VALUES FROM EACH FEATURE MAP:

Models	**PLCC**	**KLD**	**AUC**
C_1 (C_b COLOR) FEATURE MAP	0.818	0.046	0.608
C_2 (C_r COLOR) FEATURE MAP	0.960	0.040	0.606
L(LUMINANCE)FEATURE MAP	0.807	0.041	0.604
T (TEXTURE) FEATURE MAP	0.818	0.050	0.608
D (DEPTH) FEATURE MAP	0.603	0.449	0.605
FINAL SALIENCY MAP	0.517	0.111	0.604
HMM BASED SALIENCY MAP	0.823	0.022	0.609

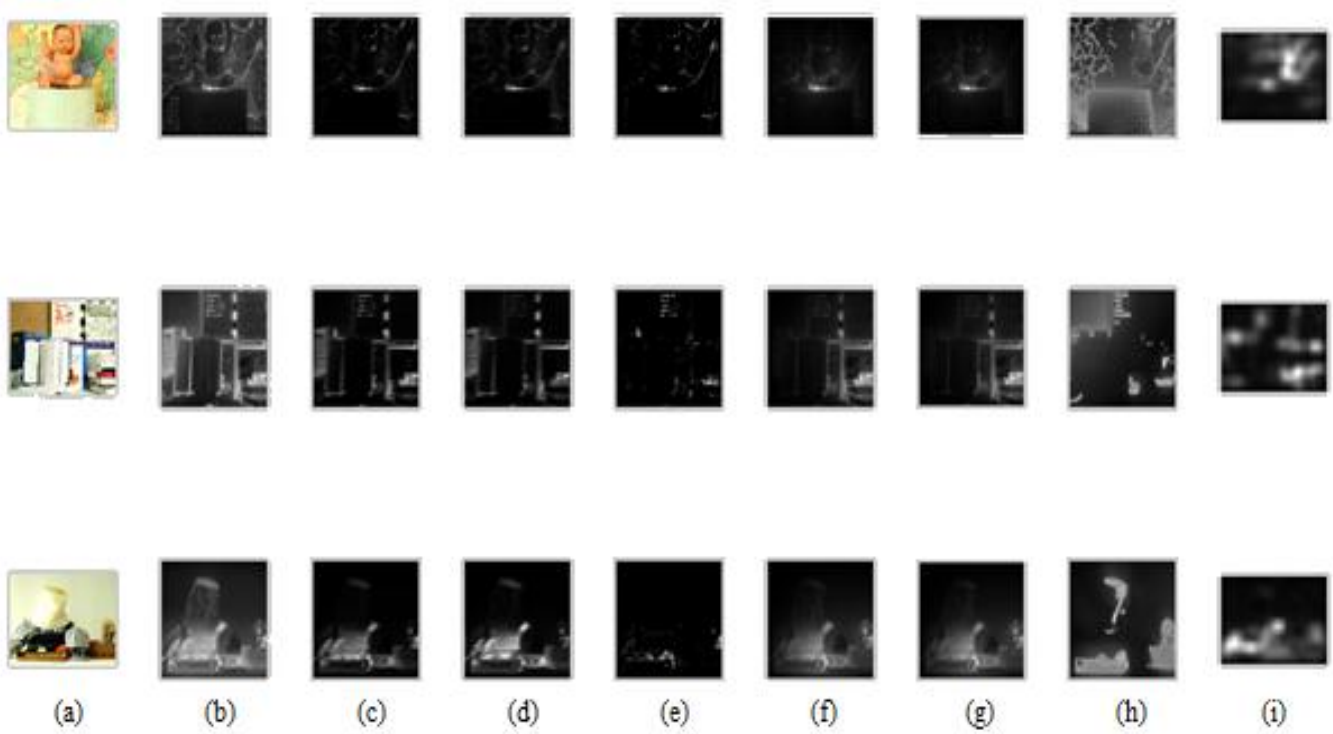

Fig. 5: Visual comparison of Saliency estimation from different features: (a) input image; (b) color feature map from Cb component; (c) color feature map from Cr component; (d) luminance feature map; (e) texture feature map; (f) depth map; (g) final saliency map; (h) ground truth map.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

ROC CURVE-I

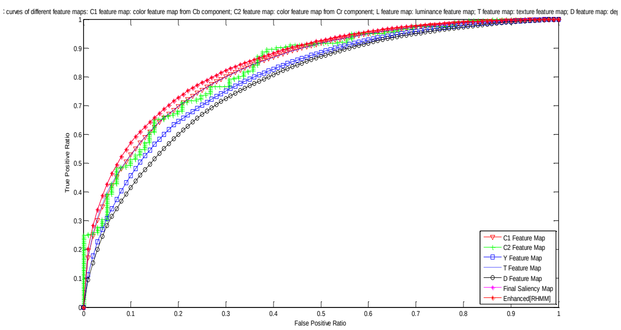

Fig. 6: The ROC curves of different feature maps: C1 feature map: color feature map from Cb component; C2 feature map: color feature map from Cr component; L feature map: luminance feature map; T feature map: texture feature map; D feature map: depth feature map.

ROC CURVE-II

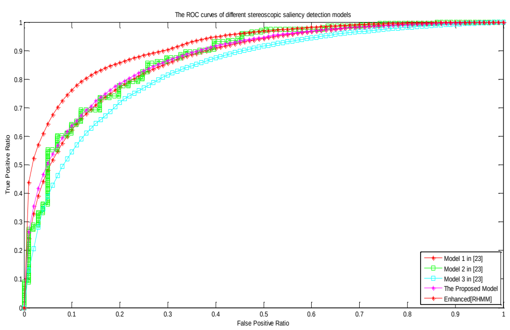

Fig.7: the ROC curves of different stereoscopic saliency detection models.

TABLE-II

COMPARISON RESULTS OF PLCC, KLD AND AUC VALUES FROM DIFFERENT STEREOSCOPIC-3D SALIENCY DETECTION MODELS; * INDICATES THAT IT IS SIGNIFICANTLY DIFFERENT FROM THE PERFORMANCE OF THE PROPOSED MODEL

Models	**PLCC**	**KLD**	**AUC**
Model 1 in [11]	0.356*	0.704*	0.656*
Model 2 in [11]	0.424*	0.617*	0.675*
Model 3 in [11]	0.410*	0.605*	0.670*
The Enhanced Model	0.928	0.204	0.604

The TABLE-II shows the efficiency of the enhanced model in detecting saliency, when compared to the models that existed in the past. The ROC CURVE-II also depicts the same.

Here the performance of obtained feature maps is compared in this experiment. The quantitative values obtained from comparison are provided by TABLE-I reflecting the performance of feature maps, where depth feature map shows better performance, followed by luminance feature map which is in turn followed by color features C_1 and C_2, in terms of increasing order of performance. The PLCC, AUC and KLD values in the TABLE-I shows that, the new fusion process used in this model resulting in a final saliency map achieved better performance in estimating saliency and is proved to be far better when compared to the feature maps from which it is obtained. Comparing the enhanced saliency detection method with the existing ones: are adopted for enhancing the saliency map. The results from experiments based on stereoscopic image databases shows the better performance of this enhanced saliency detection model.

5. CONCLUSION

The study introduces an enhanced model for saliency detection in stereoscopic images. For representing the energy of small image patches, we extract the luminance, color, texture and depth features from the coefficients of DCT. A Gaussian model is employed in the process of feature contrast calculation and, additionally a robust Hidden Markov Model (HMM) is also considered for better saliency estimation. A special fusion process is implemented to combine all the feature maps to obtain the original saliency map. In addition, the center bias factor and the human visual acuity which are the HVS chief characteristics.

REFERENCES

[1] Yuming Fang, *Member*, IEEE, Junle Wang, Manish Narwaria, Patrick Le Callet, *Member*, IEEE, and Weisi Lin, *Senior Member*, "Saliency detection for stereoscopic images", IEEE Trans. Image Proc. Vol. 23, NO. 6, June 2014.

[2] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis,"*IEEE Trans. Pattern Anal. Mach. Intell.*, vol.20, no.11, pp. 1254–1259, Nov. 1998.

[3] Y.Fang, Z.Chen, W. Lin, and C.-W. Lin, "Saliency detection in the compressed domain for adaptive image retargeting," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3888–3901, Sep. 2012

[4] Gopalakrishnan, Y. Hu, and D. Rajan, "Salient region detection bymodeling distributions of color and orientation," *IEEE Trans. Multimedia*,vol. 11, no. 5, pp. 892–905, Aug.2009.

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

[6] O. Le Meur, P. Le Callet, and D. Barba, "A coherent computational approach to model the bottom-up visual attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 802–817, May 2006.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

[7] N. Bruce and J. Tsotsos, "An attentional framework for stereo vision," in *Proc. 2nd IEEE Canadian Conf. Comput. Robot Vis.*, May 2005, pp. 88–95.

[8] Y. Zhang, G. Jiang, M. Yu, and K. Chen, "Stereoscopic visual attention model for 3d video," in *Proc. 16th Int. Conf. Adv. Multimedia Model.*, 2010, pp. 314–324.

[9] C. Chamaret, S. Godeffroy, P. Lopez, and O. Le Meur, "Adaptive 3D rendering based on region-of-interest," *Proc. SPIE*, vol. 7524, Stereoscopic Displays and Applications XXI, 75240V, Feb. 2010.

[10] N. Ouerhani and H. Hugli, "Computing visual attention from scene depth," in *Proc. IEEE 15th Int. Conf. Pattern Recognit.*, Sep. 2000,pp. 375–378.

[11] J. Wang, M. Perreira Da Silva, P. Le Callet, and V. Ricordel, "Computational model of stereoscopic 3D visual saliency," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2151–2165, Jun. 2013.

[12] A. Ciptadi, T. Hermans, and J. M. Rehg, "An in depth view of saliency,"in *Proc. BMVC*, 2013, pp. 112.1–112.11.

[13] W. S. Geisler and J. S. Perry, "A real-time foveated multisolution system for low-bandwidth video communication," *Proc. SPIE*, vol. 3299, Human Vision and Electronic Imaging III, 294, Jul. 1998.

[14] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Proc. IEEE Int. Conf. CVPR*, Jun. 2007, pp.18.

[15] C. Theoharatos, V. K. Pothos, N. A. Laskaris, G. Economou, and S. Fotopoulos, "Multivariate image similarity in the compressed domain using statistical graph matching," *Pattern Recognit.*, vol. 39, no. 10, pp. 1892–1904, 2006.