ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

Performance Analysis of Saliency Detection for Stereoscopic Images

Spandana Saggurthi[1], Rakesh.Y[2]

P.G Student, Department of ECE, SRKIT, Enikepadu, Vijayawada, India[1]

Assistant Professor, Department of ECE, SRKIT, Enikepadu, Vijayawada, India[2]

ABSTRACT: Visual attention is a mechanism to derive possible location of object or regions from natural scenes. Saliency detection technique has become a valuable tool in image processing because of many applications.Here we present methodologies based on the performance analysis of saliency detectionfor Stereoscopic Images based on MultiScale Retinex (MSR) and superpixel segmentation.The proposed method follows two steps Firstly, MSRor Superpixel Segmentation is used to enhances 2d image and the depth feature is considered in saliency detection for stereoscopic image. And next four features namely color,luminance,texture and depth are extracted by using discrete cosine transform coefficients for feature contrast calculation. Then, a fusion method is used to combine the feature mapsto get the final saliency map of stereoscopic images. Experimental results shows the performance analysis of saliency detection for stereoscopic images for the two proposed methodologies.

KEYWORDS: Stereoscopic image, Multi-Scale Retinex(MSR), Superpixel Segmentation,visual attention,saliency detection.

I. INTRODUCTION

Visual attention is an important part in the Human Visual System for visual information processing. Visual attention would selectively process the important part by filtering out others to reduce the complexity of scene analysis. These important visual information is also termed as salient regions in natural images. The visual attention applications such as visual retargeting [2],visual quality assessment [3],[4],visual coding[5],etc. The applications of stereoscopic display for 3D multimedia such as 3D video coding[6],3D visual quality assessment [7],[8],3D rendering [9],etc.Standard image enhancement techniques modify the image by using techniques such as histogram equalization, specification etc [11] so that the enhanced image is more pleasing to the visual system of the user than the original image. There is a difference between the way our visual system perceives a scene when observed directly and in the way a digital camera captures the scene. Our eyes can perceive the color of an object irrespective of the illuminant source. But the color of the captured image depends on the lighting conditions at the scene. Our aim is to enhance the quality of the image as to how a human being would have perceived the scene. The depth factor has to be taken into account for saliency detection of 3D images. The depth perception is achieved bybinocular depth cues are merged together with others in an adaptive way based on viewing space conditions the change of depth perception largely influences the human viewing behaviour [10].

The input 2D image is modified by the MSR algorithm and super pixel segmentation to get a better image for visual attention than the original image.The inputimage and depth map are divided into small image patches. Color, luminance and texture features are extracted based on DCT coefficients of each image patch from the modified input image, while depth feature is extracted based on DCT coefficients of each image patch in the depth map. Feature contrast is calculated based on center-surround feature difference, weighted by a Gaussian model of spatial distances between image patches for the consideration of local and global contrast. A fusion method is designed to combine the feature maps to obtain the final saliency map for 3D images. Additionally, inspired by the viewing influence from centre bias and the property of human visual acuity in the HVS, we propose to incorporate the centre bias factor and human visual acuity into the proposed model to enhance the saliency map. The Centre-Bias Map (CBM) calculated based on centre bias factor and a statistical model of human visual sensitivity is adopted to enhance the saliency map for obtaining the final saliency map of 3D images. Existing 3D saliency detection models usually adopt depth information to weight the traditional 2D saliency map combine the depth saliency map and the traditional 2D saliency

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

map simply to obtain the saliency map of 3D images. Different from these existing methods, the proposed model adopts the low-level features of color, luminance, texture and depth for saliency calculation in a whole framework and designs a novel fusion method to obtain the saliency map from feature maps.

The remaining part of the paper is as follows section Section II introduces the related work. In Section III, the proposed model is described in detail. Section IV provides the experimental results. The final section concludes the paper.

II. RELATED WORK

In this paper, we propose a Multi Scale Retinex(MSR) and Superpixel Segmentation for saliency detection of 3D images.

Multiscale Retinex:

The Retinex is a human perception-based image processing algorithm provides color constancy and dynamic range compression.Retinex improves visual rendering of an image when lighting conditions are not good. While our eye can see colors correctly when light is low, cameras and video cams can't manage this well. The MSRCR (MultiScale Retinex with Color Restoration) algorithm, which is at the root of the Retinex filter, is inspired by the eye biological mechanisms to adapt itself to these conditions. Retinex stands for Retina + cortex. Besides digital photography, Retinex algorithm is used to make the information in astronomical photos visible and detect, in medicine, poorly visible structures in X-rays or scanners.

Superpixel Segmentation:

Superpixels (image segments) can provide powerful grouping cues to guide segmentation, where superpixels can be collected easily by segmenting the image using any reasonable existing segmentation algorithms. Generated by different algorithms with varying parameters, superpixels can capture diverse and multi-scale visual patterns of a natural image. Successful integration of the cues from a large multitude of superpixels presents a promising yet not fully explored direction. In this paper, we propose a novel segmentation framework based on bipartite graph partitioning, which is able to aggregate multi-layer superpixels in a principled and veryeffective manner.

Our goal is to use the superpixels as basic units in the salient object detection and in the object recognitiontask in our future work. With respect to this goal the desired method is expected to fulfil the criteriaSpatial coherence of clusters in image space.The boundary of the segmented superpixel should follow the boundary of regularly sized and regularlydistributed rectangular regions as far as no saliency edge are in the neighbourhood.Local spatial maxima and local minima in image, whose area is small compared with the desired sizeof the segmented superpixel (irrespective the corresponding intensity value of this local extreme) will beregarded as irrelevant and will have no contribution to the segmentation.Our approach is based on morphological processing and use of 8-connectivity. These methods are timeefficient and guarantee a spatial coherence. The crucial point is a question which edge is a saliency edge andwhich edge should be ignored in the process of superpixel boundary For this decision we have constructed and enhanced image which has been done by removing of all regional local intensity extreme, hence removingall edges with local significance. The remaining edges are global and will be accepted as saliency edges.

III. THE PROPOSED MODEL

The framework of the proposed model is shown in Fig1. Firstly, the input image is modified by using MSR or superpixel segmentation and then thecolor, luminance, texture, and depth features are extracted from the input stereoscopic image. Based on these features, the feature contrast is calculated for the feature map calculation. A fusion method is designed to combine the feature maps into the saliency map. The input image is divided into small image patches and then the DCT coefficients are adopted to represent the energy for each image patch. The used image patch size is also the same as DCT block size in JPEG compressed images. The input RGB image is converted to YcbCrcolor space due to its perceptual property. In YcbCrcolor space, the Y component represents the luminance information, while Cb and Cr are two color-opponent components. For the DCT coefficients, DC coefficients represent the average energy over all pixels in the image patch, while AC coefficients represent the detailed frequency properties of the image patch.Thus, we use the DC coefficient of Y component to represent the luminance feature for the image patch as $L=Y_{DC}$ (Y_{DC} is the DC coefficient of Y component), while the DC coefficients of Cb and Cr components are adopted to represent the color features as $C_1=C\ b_{DC}$ and $C_2=Cr_{DC}$ ($C\ b_{DC}$ and Cr_{DC} are the DC coefficients from Cb and Cr components respectively).Since the Cr and Cb components mainly include the color information and little texture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

information is included in these two channels, we use AC coefficients from only Y component to represent the texture feature of the image patch.

In DCT block, most of the energy is included in the first several low-frequency coefficients in the left-upper corner of the DCT block. As there is little energy with the high-frequency coefficients in the right-bottom corner of the DCT block, we just use several first AC coefficients to represent the texture feature of image patches.

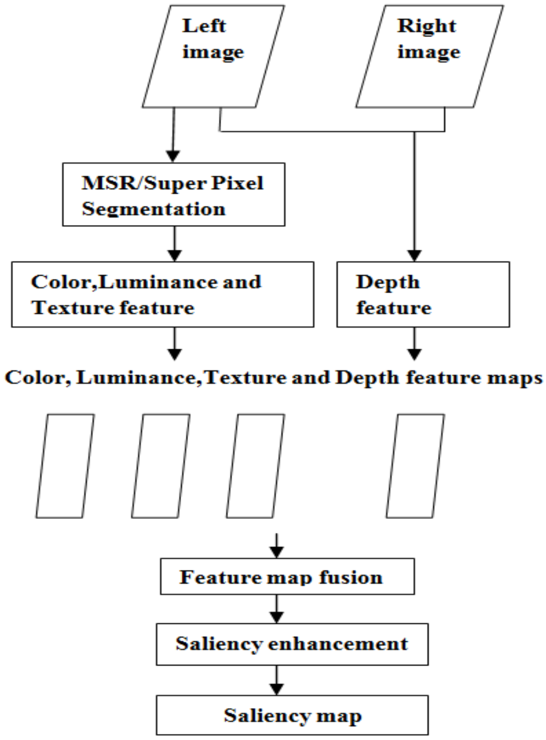

Fig1: frame work of proposed model

The existing study demonstrates that the first 9 low-frequency AC coefficients in zig-zag scanning can represent most energy for the detailed frequency information in one 8 × 8 image patch. Based on the study, we use the first 9 low-frequency AC coefficients to represent the texture feature for each image patch as $T= \{Y_{AC1},Y_{AC2}, . . . ,Y_{AC9}\}$.depth map provides the information of the perceived depth for the scene. In a stereoscopic display system, depth information is usually represented by a disparity mapwhich shows the parallax of each pixel between the left-view and the right-view images. The disparity is usually measured in unit of pixels for display systems.The five features of color, luminance, texture and depth (L,C_1,C_2,T,D) for the input stereoscopic image are extracted by using DCT. The feature maps are calculated[1] , saliency estimation from feature map fusion is estimated[1]and then saliency enhancement is done for the obtained feature for final saliency map[1]

IV. EXPERIMENT EVALUATION

In this section, we conducted the experiments to demonstrate the performance of the proposed 3D saliency detection model. We first present the evaluation of color, luminance, texture and depth feature maps,2d and 3d saliency maps.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

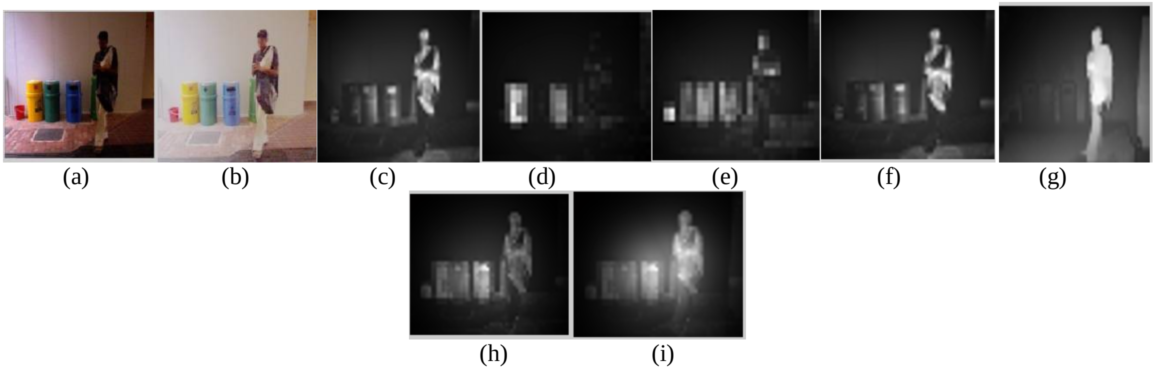

Fig2: visual samples of saliency detection based on MSR (a) original image (b) Retinex Image (c)color feature map from cb component (d) color feature map from cr component (e) Luminance feature map (f) texture feature map (g) depth feature map (h) 2d saliency map (i) 3d saliency map.

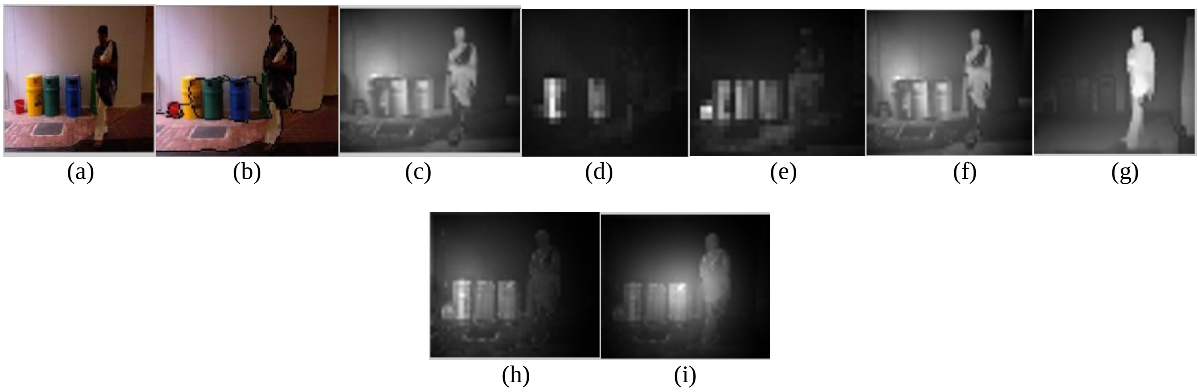

Fig3: visual samples of saliency detection based on Super pixel segmentation :(a) original image (b) Super pixel segmentation Image (c)color feature map from cb component (d) color feature map from cr component (e) Luminance feature map (f) texture feature map (g) depth feature map (h) 2d saliency map (i) 3d saliency map.

Table1:Measurements of methodologies

Comparison results of segment error,log spectral distance, and quality for the proposed two different methodologies

METHODOLOGIES	SEGMENT ERROR	LOG SPECTRAL DISTANCE	QUALITY
MSR+SALIENCY DETECTION	4.496*e^-5	1.010*e^-5	0.0838
SUPER PIXEL SEG+SALIENCY DETECTION	4.455*e^-5	1.0056*e^-5	0.0836

V. CONCLUSION

We proposed the saliency detection based on (MSR) and superpixel segmentation for Stereoscopic Images. To improve the image quality, MSR or Superpixel Segmentation is applied to the input image. The saliency is estimated based on the energy contrast weighted by a Gaussian model of spatial distances between image patches for the consideration of both local and global contrast. A fusion method is designed to combine the feature maps for the final

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

saliency map. Experimental results show the performance of the saliency detection for stereoscopic images based on the MSR and superpixel segmentation.

REFERENCES

[1]Yuming Fang, Junle Wang, Manish Narwaria, Patrick Le Callet, and Weisi Lin, "Saliency detection for stereoscopic images**",** IEEE Transaction on Image Processing,Vol. 23, No. 6, 2625-2636, 2014.**[2]**Y. Fang, Z. Chen, W. Lin, and C.-W. Lin, "Saliency detection in the compressed domain for adaptive image retargeting," IEEE Trans. ImageProcess., vol. 21, no. 9, pp. 3888–3901, Sep. 2012.
[3]Z. Lu, W. Lin, X. Yang, E. Ong, and S. Yao, "Modeling visual attention's modulatory aftereffects on visual sensitivity and quality evaluation," IEEE Trans. Image Process., vol. 14, no. 11, pp. 1928–1942, Nov. 2005.
[4]W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey,"J. Visual Commun. Image Representation, vol. 22, no. 4, pp. 297–312,2011.
[5]C. Guo and L. Zhang, "A novel multi-resolution spatiotemporal saliency detection model and its applications in image and video compression," IEEE Trans. Image Process., vol. 19, no. 1, pp. 185–198, Jan. 2010.
[6]F. Shao, G. Jiang, M. Yu, K. Chen, and Y.-S.Ho, "Asymmetric coding of multi-view video plus depth based 3D video for view rendering," IEEE Trans. Multimedia, vol. 14, no. 1, pp. 157–167, Feb. 2012.
[7]F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," IEEE Trans. Image Process., vol. 22, no. 5, pp. 1940–1953, May 2013.
[8]Q. Huynh-Thu, M. Barkowsky, and P. Le Callet, "The importance of visual attention in improving the 3D-TV viewing experience: Overview andnew perspectives," IEEE Trans. Broadcast.,vol.57,no.2,pp421–431, Jun. 2011.
[9]C. Chamaret, S. Godeffroy, P. Lopez, and O. Le Meur, "Adaptive 3D rendering based on region-of-interest," Proc. SPIE, vol. 7524, Stereo-scopic Displays and Applications XXI, 75240V, Feb. 2010.
[10]J. Hakkinen, T. Kawai, J. Takatalo, R. Mitsuya, and G. Nyman, "What do people look at when they watch stereoscopic movies?"Proc. SPIE, vol. 7524, Stereoscopic Displays and Applications XXI, 75240E, Feb. 2010.
[11]R. C. Gonzalez, R. E. Woods, "Digital Image Processing," Third Edition, Pearson Publications**.**

BIOGRAPHY

Saggurthi Spandana received the B.Tech degree in electronics and communication engineering from LakiReddy Bali Reddy College of engineering Mylavaram, and pursuing M.Tech in SRK Institute of Technology, Vijayawada, Andhra Pradesh, India. Research interest are Visual Attention Modeling and Object Detection/Recognition

Rakesh.Y working as Assistant prof at SRKIT in the Dept. of ECE completed his M Tech at Vignan Engineering College and pursuing his PHD at Nagarjuna University. Research Interests are Image Processing and Signal Processing.

