

Visually Salient Region Detection in Omnidirectional Images Using Wavelet Textural Feature Map

Manisha Mane^{1,A,B}, Anand Bhaskar^{2,A}

^A School of Engineering, Sir Padampat Singhanian University, Bhatewar, Udaipur-313601, Rajasthan, India

^B Shah and Anchor Kutchhi Engineering College, Chembur, Mumbai-88, Maharashtra, India

¹ ORCID: 0009-0005-4832-3874, manishamane.sakec@gmail.com

² ORCID: 0000-0002-6255-328X, anand.bhaskar@spsu.ac.in

Abstract

Salient object detection is a crucial aspect of computer vision that involves identifying the most prominent area in a 2D image. However, predicting salient regions in omnidirectional images can be challenging due to their circular field of view. To facilitate saliency detection, pre and post-processing are required, which involves converting the image into an equirectangular projection (ERP). In this study, we propose a detailed approach to saliency detection for omnidirectional images using the wavelet domain. Our proposed model utilizes a 2-D wavelet transform to decompose and reconstruct images in the CIELAB space. The texture channel map is then calculated, followed by the feature map, where salient regions are marked using Gaussian filtering and entropy. Our experimental results demonstrate that this method is highly effective for detecting salient objects in omnidirectional images.

Keywords: ERP, Omnidirectional image, Saliency detection, Salient object map, Saliency model, Visual saliency, 2-D wavelet transform.

1. Introduction

Saliency detection, also known as salient-object detection, involves identifying significant objects within an image. This process enables a system to quickly recognize important regions, similar to how the human visual system operates. By focusing on key objects within an image, a computer can better understand the scene's overall importance, making it particularly useful for capturing context information or detecting shapes. Salient region detection methods can automatically capture important image features such as color, edges, and boundaries [1]. The field of computer vision has explored the process of saliency detection and its importance for a variety of applications, including object recognition, scene perception, image segmentation, image retrieval, adaptive image or video compression, and medical imaging [2][3].

In general, effective saliency identification methods should prioritize accurate object detection, computational efficiency, and high resolution [4]. Researchers have established a range of standards for saliency prediction, categorized by bottom-up and top-down approaches. Over the last few years, much work has been conducted on saliency detection in both the spatial and frequency domains. A comprehensive review of research conducted in this field is provided in [2]

Feature extraction in the spatial domain is performed using pixel coefficients. Several categories of saliency detection are based on local and global contrast, with detection relying on center-prior, backgroundness prior, and objectness prior. Many saliency methods have been proposed using color spatial distribution, center surround histogram, and multi-scale con-

trast. However, the spatial domain has limitations; the salient object is within a rectangular box and its boundary cannot be detected. Additionally, spatial detection methods are computationally expensive and require appropriate parameter selection [5]. Some methods cannot clearly differentiate the image from its background [6], likely due to the presence of high-frequency content in the image [7]. In some cases, salient region boundaries are highlighted, but the salient region is not consistently featured. All of these limitations arise from the presence of inappropriate frequency content in the image. To address these issues, the frequency domain is often preferred [8].

The process of saliency detection in the frequency domain involves several steps. First, the frequency spectrum of the input image is obtained by taking its frequency response to eliminate the image background. Next, image enhancement techniques are applied to highlight the salient regions, and a gray level saliency map is generated by taking the inverse frequency transform. The first saliency detection method in the frequency domain was proposed by Hou et al. [9], who discovered that statistical singularities in the amplitude spectrum of the Fourier transform (FT) were responsible for salient objects, and that the average log amplitude spectrum contained redundancies. Guo et. al. [10], then constructed a 2D signal using phase spectrum to obtain a saliency map, and combined the color and intensity features of the image into a quaternion Fourier transform (QFT) to suppress the non-salient regions. The saliency map was generated based on a minimum entropy criterion by convolving the amplitude spectrum with a low-pass Gaussian kernel at different scales. Fourier Transform is not suitable for analyzing natural images that are non-stationary signals, and it does not provide simultaneous spatial and frequency resolution of signals. Therefore, the short-time Fourier transform (STFT) is used for frequency analysis [11], where the non-stationary image signal is segmented into narrow spatial intervals, considered stationary, and then Fourier transform is applied. The limitation of STFT is non-uniform resolution, where high frequency resolution is required in the spatial domain and low frequency resolution is important in the frequency domain [12]. STFT also leads to high computational complexity. To address these shortcomings, new saliency detection methods based on discrete wavelet transform (DWT) have been proposed [13][14][15][16]. DWT enables simultaneous analysis of spatial and frequency domains, and its basis functions vary in frequency and time for a limited duration, making multi-scale analysis possible. The saliency map is obtained by taking the inverse discrete wavelet transform (IDWT) for each color sub-band, where weights are derived from the high-pass wavelet coefficients of each level [17][11]. However, these methods can only predict salient points, not salient regions. In advanced methods, feature maps are obtained by taking DWT of the image at each decomposition level, and the IDWT is taken after setting low-pass coefficients of that level to 0. Local contrast is obtained by considering linear combinations of the feature maps, and the computation of global distinctiveness feature maps is based on the normal distribution of the features [14][18][19].

Omnidirectional Images also known as 360-degree Images are used in wide range of applications such as robotic vision, surveillance etc. due to their large field of view (FOV). The 360 degree cameras acquires the image which is equivalent to the placing of multiple conventional cameras, due to their narrow field of view. Also the optical axes for each image obtained by conventional cameras is different which makes it unsuitable for many applications which demands hights FOV. The process of visual exploration for omnidirectional images (ODI) differs from that of conventional 2D images because ODIs provide a wider perspective. Detecting visual saliency in such images is not as straightforward as in conventional images since the viewer may be interested in only a specific part of the ODI rather than the entire image. Therefore, conventional 2D models for saliency detection cannot be applied directly to ODI images. Pre- and post-processing steps are required for these images, which involve unwrapping the image and converting it into an Equirectangular Projection (ERP) image to enable saliency detection [20][21][22][23][24].

We present a saliency prediction method for omnidirectional images in this paper, which employs wavelet transform. The experiments are conducted using an ERP image, and the results are calculated accordingly. The paper is structured as follows: Section 2 explains the process of converting omnidirectional images to ERP images; Section 3 details the proposed saliency detection technique that employs DWT and IDWT. Section 4 investigates the impact of Gaussian filtering and entropy calculations. Section 5 provides a discussion of the experimental results obtained from various techniques, along with a comparison with other existing techniques. Finally, Section 6 concludes the paper.

2. Conversion of OD image into ERP image

Omnidirectional images contain information in circular form, which requires a different approach to visual exploration than traditional 2D images. The expanded field of view provided by omnidirectional images also presents challenges for visual saliency detection. Unlike 2D images, a viewer may not be interested in the entire image but only a particular region, leading to the need for unwrapping omnidirectional images. Three major image unwrapping techniques are available: cube projection, sphere projection, and the Equirectangular projection method (ERP). Of these methods, the most preferable one is the ERP method, which converts omnidirectional images into rectangular images, as depicted in Figure 1[22][23][24].



Figure 1 (a): Omnidirectional Image (b): Unwrapped equirectangular projection (ERP) Image

Due to the large size and panoramic view of the ERP image, it is not appropriate for analysis using DWT. Thus, it is necessary to segment the ERP image for use in pre-processing saliency detection algorithms. The ERP image is segmented into several sub-images. In post-processing, an image stitching algorithm can be used to obtain the saliency detection of the ERP image. The segmented image obtained from the ERP image is depicted in Fig. 2.



Figure 2: (a) ERP Image, (b) Segmented ERP Image 1(I_{s1}), (c) Segmented ERP Image 2(I_{s2}).

To detect saliency in the segmented ERP images, DWT is applied at various levels. Wavelet decomposition is performed up to level N , and to obtain the feature map at level N , the low-pass coefficients of that level are set to zero, and IDWT is applied [25][17]. Since the last decomposition level contains all the detailed coefficients of the previous levels, IDWT does not need to be applied at each decomposition level. Additionally, generating feature maps at each decomposition level can increase the computational complexity of the method [26].

The method proposed in this paper employs DWT-based textural features of different color channels to detect salient regions in images. Before detecting saliency from an image, the first step is to select an appropriate color space for image analysis. The commonly used techniques for color space analysis are RGB, YCbCr and CIELAB.

The RGB color picture is sensitive to changes in luminance color, and any color space can be obtained using linear or non-linear combinations of RGB. However, it is device-dependent, resulting in variations in the signal on different devices, and its chrominance and luminance components are mixed, making it less preferable for color analysis [27]. On the other hand, YCbCr is the most widely used color space, especially in real-time digital video processing and information handling. Y represents luminance, while Cb and Cr represent chrominance components. It is computed using a non-linear combination of RGB [28][7].

The most commonly used perceptually uniform color space is CIELAB where the non-linear combinations are used to obtain colors and is useful when exact color definition is needed. It is independent of input colors and also includes more colors than any other color space [29][30]. In the proposed algorithm, the input SERP image is first converted into CIELAB color space, which uses L, a, and b to indicate luminance, red/green, and blue/yellow channels, respectively [1][5][28]. Fig.3 illustrates the complete process of the proposed method. Firstly, feature maps for each channel of the CIELAB color space are obtained through DWT decomposition. The channel feature maps are generated based on the entropy value, and the final saliency map is obtained after low pass filtering of the image feature map [11]. After applying the SERP images to the CIELAB color space, separate L band, a band, and b band images are obtained, as illustrated in Fig. 4. Once these images are obtained, the channel feature map is computed by extracting the features of the lab-separated image. To obtain all the coefficients of the image at N different decomposition levels, the discrete wavelet transform is used.

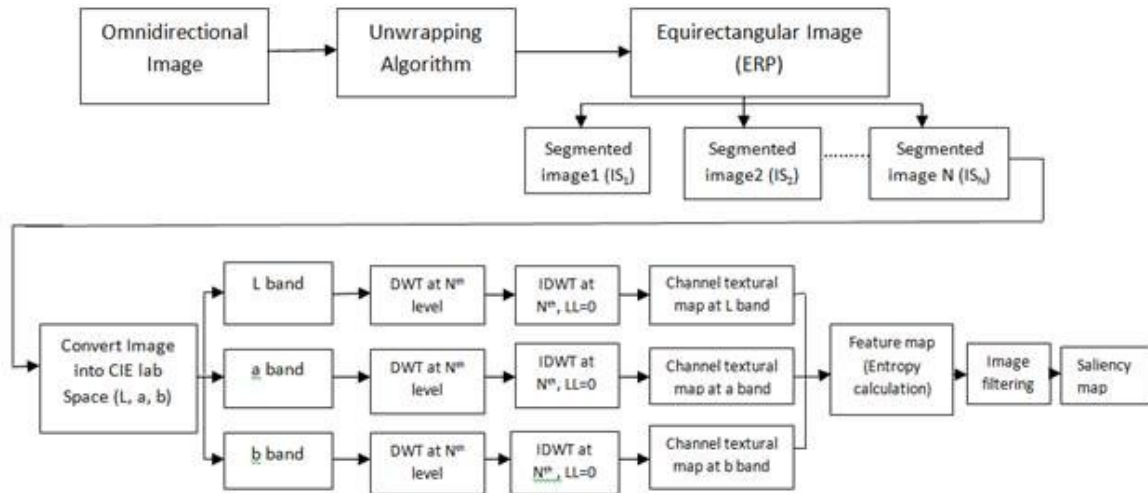


Figure3: Process flow of the algorithm



SERP Image



L band Image



aband Image



bband Image

Figure 4: CIELAB color space images (L, a, b images respectively)

3. Saliency detection using DWT and IDWT

The transformation of images from the spatial domain to the frequency domain is accomplished using the Wavelet transform. This method yields four types of coefficients, namely LL, LH, HL, and HH, which describe the high and low features of the images depicted in Fig. 5. To obtain features at high-resolution levels, multiple layers of decomposition are required [30]. The inverse wavelet transform of the low-frequency feature map is used to obtain the local feature saliency map. In the proposed method, the Discrete Wavelet Transform (DWT) is utilized to extract textural features of the image at various scales $Se\{1, \dots, N\}$, where N is the highest decomposition level of DWT. The DWT is carried out independently on the L, a, and b channels of the SERP image [4].

$$DWT(I^c) = \{A_N^c, \{H_s^c, V_s^c, D_s^c\}; S \in \{1, \dots, N\}\}$$

where DWT is the discrete wavelet transform of the image with c color channels (I^c) and $c \in \{L, a, b\}$, represents the color channels of the input image. The base level N approximation coefficient is denoted by A_N^c , also known as the LL band. The detailed coefficients of the image are represented by H_s^c, V_s^c, D_s^c , whereby H_s^c is the horizontal (LH) coefficient, V_s^c is the vertical (HL) coefficient matrix, and D_s^c is the diagonal (HH) coefficient matrix of the channel cat level s , respectively [11]. The images with level 1 decomposition coefficients of each color space of the input image are shown in Fig. 6.

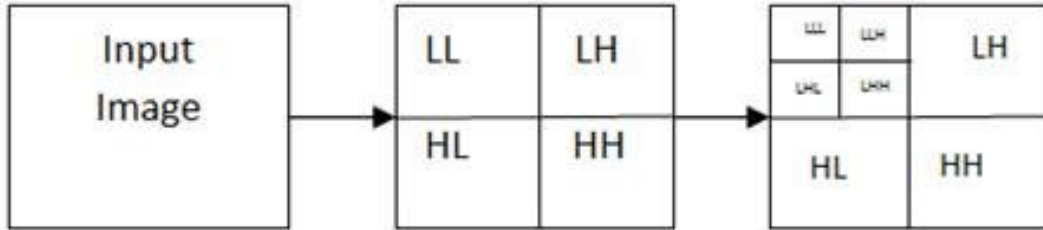


Figure 5: Discrete wavelet transform (DWT) decomposition

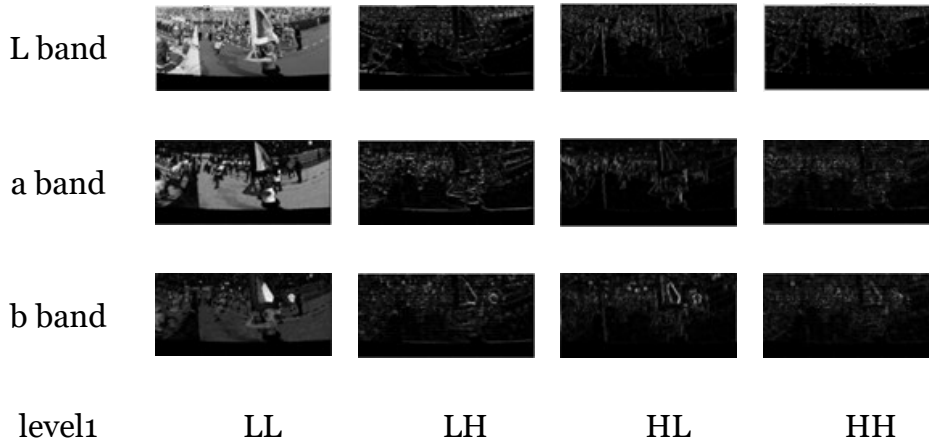


Figure 6: DWT decomposition coefficients of color space input Image

To extract textural details from the DWT decomposition of the image, a higher level of decomposition is employed. Fig. 7 illustrates the second-level decomposition of the b band image



Figure 7: 2nd level decomposition of b band image (a)LL band, (b)LH band, (c) HL band, (d) HH band

To obtain the textural feature map, the coefficients A_N^c at the highest decomposition level N are set to 0, and a 2D image is reconstructed by applying the IDWT that can be represented as follows:

$$T_{MAP}^c = IDWT(A_N^c = 0, \{H_S^c, V_S^c, D_S^c\}; S \in \{1, \dots, N\},$$

here the texture map of channel c is T_{MAP}^c , which is acquired by utilizing the inverse DWT. The reconstructed images for the L, a, b band images, and our proposed method image, are displayed in Fig. 8. In general, salient region is location of salient objects in the image having a high pixel intensity level associated with them. Salient objects tend to be centrally located and thus, weights are calculated without considering their proximity to the image borders. The channel feature map of the image is generated by enhancing the features of T_{MAP}^c . To achieve this, the cluster of pixels with high intensity values is strengthened at location (i, j) of the image using $T_{MAP}(i, j)$, which is known as the feature map (F_{MAP}). This can be expressed as follows:

$$F_{MAP}^c(i, j) = (T_{MAP}^c(i, j))^2$$

Fig. 8 illustrates the feature maps for the L, a, and b channels. Based on the images generated through the proposed algorithm, it can be concluded that an increase in the number of decomposition levels leads to an improvement in the quality of the textural feature map for $n=1, \dots, N$. As a result, the resulting feature map contains pixels that are clustered with higher gray levels compared to other pixels.

4. Effect of Gaussian filtering and entropy calculation

The image feature map is calculated by taking summation of weight of the channel c , as ω^c of the channel feature maps, which calculated as.

$$F_{MAP} = \sum_c \omega^c F_{MAC}^c$$

To calculate the weight, image entropy is used, which works as the statistical measurement tool of randomness. The channel feature map with larger entropy should be assigned a smaller weight and vice versa. Two different images with variable pixel distributions may have the same entropy value. This is the limitation of the entropy based weight determining method. To overcome this limitation entropy value can be used in a more efficient way by considering the spatial distribution of gray levels of pixels; the gray levels can be enhanced when the channel feature map is combined with low pass filtering. With the help of low pass filtering the number of gray levels will increase and hence their entropy will also vary.









Color space	L	A	B	Proposed method image
Gaussian filtered Image with Sigma 1				
Entropy	0.3291	0.5580	0.6445	0.0179
Gaussian filtered Image with Sigma 2				
Entropy	0.3031	0.5486	0.5974	4.2760e-04

Figure 8: Feature map for L,a,b band with entropy with sigma 1 and 2

Fig. 8 shows the gray level pixel distribution with entropy value and the effect of low pass filtering on the image with improved entropy value. Filter used here is a 3×3 low-pass Gaussian filter. It is observed that after the filtering the entropy value is increased. Thus entropy value is the important parameter in assigning weights and entropy is calculated as

$$\epsilon^C = H(G * FMAP^C)$$

where \mathbf{G} is a low-pass Gaussian filter and H is the entropy. The $*$ represents the 2D convolution. For the Gaussian filtering it is expected to have a high value of the standard deviation. Gaussian filter at the location (i,j) of the image is as given below and this Gaussian model is preferred to highlight center part of the image and to suppress surrounding [30][31]

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(i-i')^2 + (j-j')^2}{2\sigma^2}}$$

Here σ is standard deviation of the low-pass Gaussian filter \mathbf{G} [5] and it is calculated by number of rows(L) and columns(W) of the original image

$$\sigma = 0.02 \frac{L \times W}{2}$$

and Gaussian mask K is set to

$$\sigma = 0.25 \frac{L \times W}{2}$$

Images shows in figure 8 provide the entropy values for different standard deviation functions. Feature map for CIE lab color space with entropy with sigma 1 and 2.

The weightage of the channel feature map can be determined by its entropy value, with lower values indicating greater weight and vice versa. Our algorithm has the lowest entropy value of all, as demonstrated in Figure 8, which displays the three-color channels and their respective channel feature maps. The salient region is better represented by the channel feature map corresponding to the channel a band image, compared to the other channels L band and b band. The b channel feature map captures some details of the salient region that are not captured by the L channel feature map. The proposed method produces better results for these images.

5. Experimental results and discussion

In this section, we evaluate the performance of our algorithm using images from commonly used natural image databases, including the MSRA-10K dataset (consisting of 1000 imag-

es), the Complex Scene Saliency Detection (CCSD) dataset (containing 200 more complex background images), and the Extended Complex Scene Saliency Detection (ECSSD) dataset (containing 1000 images with different groups of the same size) [11]. The process of obtaining salient detection regions is similar to binary classification and requires a ground truth binary image (G). The obtained saliency map image (SM) is transformed into a binary map image (SMb) using an appropriate threshold value, which is selected using the indirect, adaptive threshold method[3][7]

$$TH(adpt) = \frac{2}{L \times W} \sum \sum S(i, j)$$

To evaluate the performance of the salient map, we use the ground truth value $G(i, j)$ at location (i, j) . The evaluation is both qualitative and quantitative, with Precision (P), Recall (R), and F-measure (F) - also known as F1 parameters - serving as quantitative performance measures. Additionally, we need to calculate Accuracy, which is the percentage of correctly identified pixels to the total pixels in the image. It is given by the ratio of

Accuracy = $\frac{\text{TruePositive}}{\text{TruePositive} + \text{TrueNegative}}$. The Precision performance parameter represents the positive predictive value and is given by the ratio of Precision = $\frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}$. Recall, on the other hand, is also known as sensitivity and is given by the ratio of Recall = $\frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}$.

Ideally, we want to have high precision and recall values, however, these two measures are inversely proportional, we may get good precision but bad recall or vice versa. Therefore, we combine both precision and recall into a single measure, the F-measure, which we calculate using the formula

F-Measure = $(2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$ and

$$P = \frac{\sum_i \sum_j SMb(i, j) \times G(i, j)}{\sum_x \sum_y SMb(i, j)}$$

$$R = \frac{\sum_i \sum_j SMb(i, j) \times G(i, j)}{\sum_x \sum_y G(i, j)}$$

$$F = \frac{(1 + \eta) \times P \times R}{(\eta \times P) + R}$$

In general we want excellent precision value along with high recall value it provides a way to combine both precision and recall into a single measure that to get both properties [3][8][32]. We have used the images from different data sets and the images I_{s1} and I_{s2} are SERP Image. F measure is calculated for our algorithm using given formulae.

The images are qualitatively compared in Fig. 9 and subsequently utilized to derive quantitative results by calculating the F1 parameter. Fig. 9 also displays the performance evaluation of the saliency map for images belonging to the L band, a band, b band, and the proposed method. It is evident from the results that the L band images are able to capture all the intricate details of the background and salient objects, which is suitable for images with a simple background. However, for images with a complex background, the output of this band is inadequate. On the other hand, a band and b band images produce good results for both simple and complex background images by effectively suppressing the background. Additionally, the proposed method images are obtained by combining the coefficients of the L, a, and b bands to generate more detailed images, which exhibit similar results to the a and b band images. Finally, Images I_{s1} and I_{s2} also demonstrate appropriate results for a band, b band, and proposed method images.

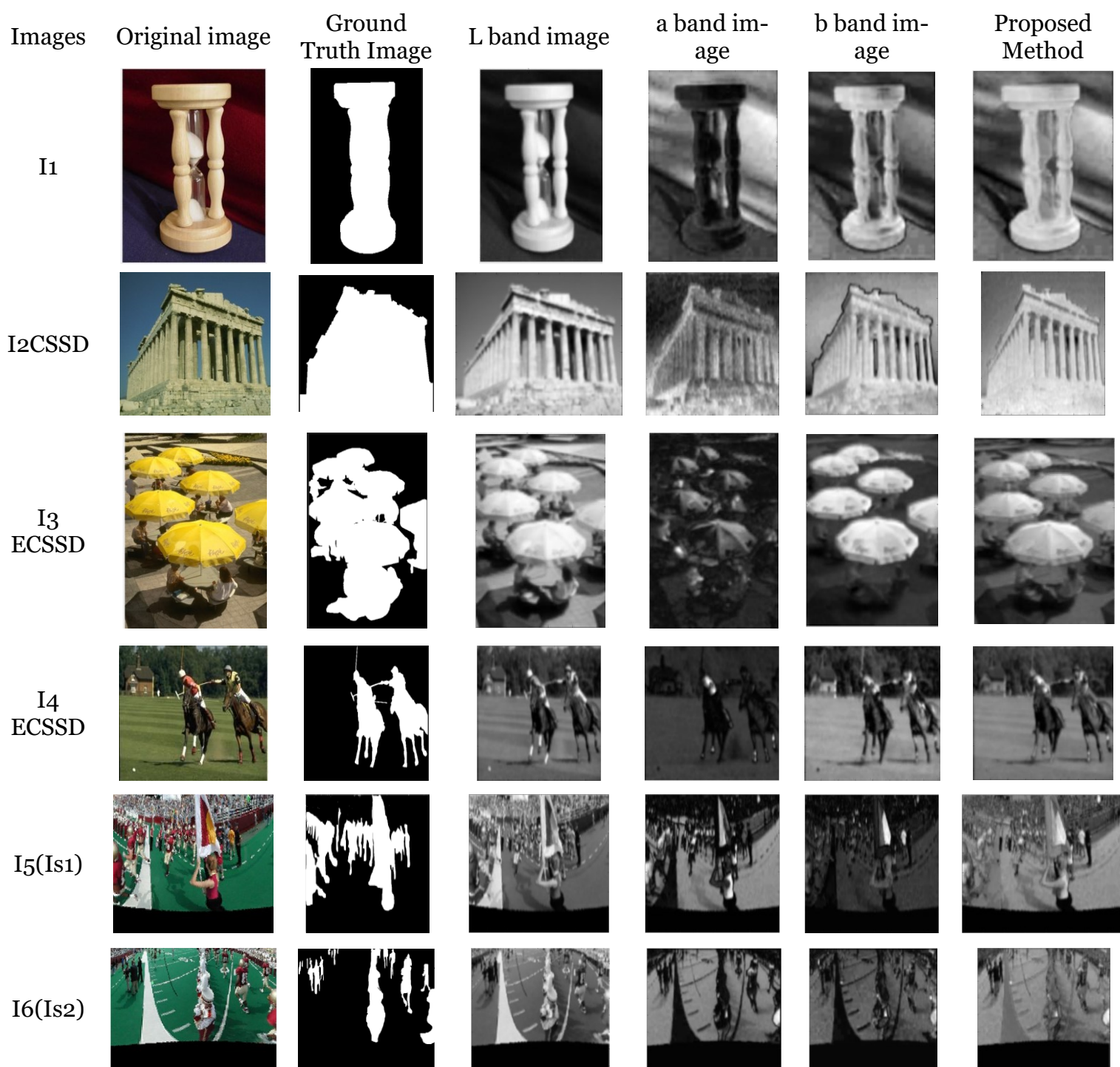


Figure 9: Image visualization of each method from left to right:(a) Input imae, (b) Ground truth image, (c) Salient map of L band image, (d) Saliency map of a band image, (e) Salient map of b band image, (f) Saliency map of proposed method image.

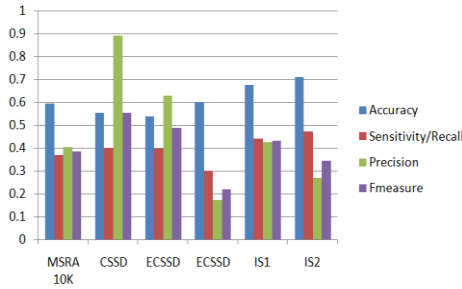
Table1:F1 parameter i) Accuracy, ii) Sensitivity, iii)Fmeasure iv) Precision

I ₁	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.5954	0.3697	0.4063	0.3871
a band	0.6287	0.2504	0.4356	0.3180
b band	0.6461	0.2917	0.4805	0.3630
Proposed Method	0.6505	0.2915	0.4908	0.3657
I ₂	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.5548	0.4031	0.8920	0.5553

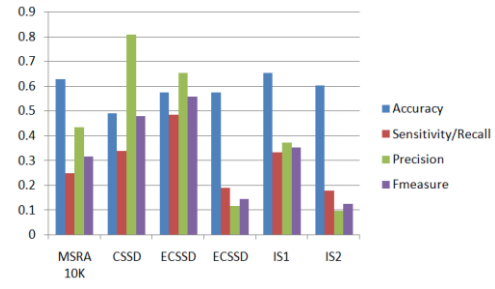
a band	0.4897	0.3403	0.8088	0.4790
b band	0.5001	0.3937	0.7682	0.5206
Proposed Method	0.3976	0.1939	0.7411	0.3074
I ₃	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.5394	0.3996	0.6285	0.4886
a band	0.5750	0.4847	0.6538	0.5566
b band	0.5490	0.3314	0.6876	0.4473
Proposed Method	0.5407	0.3550	0.6522	0.4597
I ₄	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.5998	0.3007	0.1759	0.2219
a band	0.5743	0.1897	0.1170	0.1447
b band	0.6195	0.3360	0.2004	0.2511
Proposed Method	0.5990	0.3575	0.1956	0.2529
I _{s1}	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.6751	0.4410	0.4278	0.4343
a band	0.6519	0.3342	0.3717	0.3520
b band	0.6650	0.3906	0.4044	0.3974
Proposed Method	0.6582	0.3398	0.3805	0.3590
I _{s2}	Accuracy	Sensitivity/Recall	Precision	Fmeasure
L band	0.7093	0.4736	0.2722	0.3457
a band	0.6019	0.1787	0.0986	0.1271
b band	0.6569	0.2506	0.1550	0.1915
Proposed Method	0.6567	0.1999	0.1318	0.1589

Table 1 presents a performance comparison of various types of images using the F1 parameters. The images utilized in this study are of different categories, such as I₁ and I₂, which are images with a simple background, obtained from the MSRA 10K dataset. Images I₃ and I₄, on the other hand, have a more complex background and are selected from the CSSD dataset. Additionally, images I_{s1} and I_{s2} represent the segmentation of the ERP images acquired from our algorithm. All the performance graphs are plotted for the images specified in Figure 9. Figure 10 displays the overall performance of the F1 parameters, which includes four plots. The accuracy plot provides similar results for all image types, while the sensitivity plot, the F-measure plot, and the precision plot are specific to each image type. The behavior of segmented images I_{s1} and I_{s2} is similar to that of the other image types, as indicated by the similarity in the values obtained for all these parameters. Therefore, it can be concluded that salient object detection is feasible using this method.

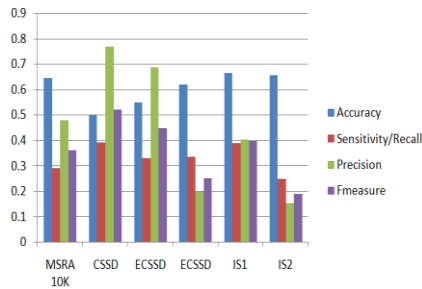
L band



a band



Bband



Proposed method

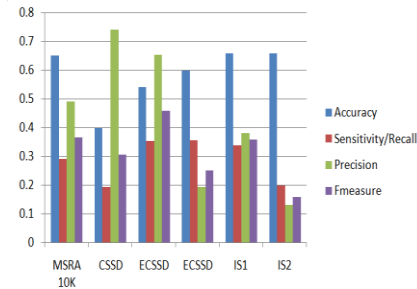


Figure 10: Performance of F1 measure.

6. Conclusion

The proposed method for detecting salient objects in omnidirectional images involves utilizing wavelet-based feature maps. First, the omnidirectional images are preprocessed and converted into segmented equirectangular projection images (SERP). Then, the CIE lab space is employed to obtain L, a, b band images for extracting textural feature maps. The textural features from SERP images are obtained using wavelet transform, where the approximation coefficients of the N^{th} level are set to zero. Calculation of feature maps is performed based on entropy and low-pass filtering, taking into account the distribution of pixels and image intensity values. Images with higher entropy values are given lower weights and vice versa. The channel feature maps are constructed by combining the textural features of the L, a, b bands. The quantitative and qualitative results reveal that the saliency detection performance of the SERP images is comparable to that of conventional images. Therefore, the proposed method is effective in identifying salient objects and is well-suited for saliency detection in omnidirectional images.

References:

1. M Banitalebi-dehkordi, M Khademi, and A Ebrahimi-moghadam, "An image quality assessment algorithm based on saliency and sparsity," *Multimed Tools Appl* **78**, pp 11507–11526, 2019
2. M Jian, K Lam, S Member, J Dong, and L Shen, "Visual-Patch-Attention-Aware Saliency Detection," *IEEE Transactions on Cybernetics* **45**, vol 45, no 8, pp 1575–1586, 2015
3. I Ullah, M Jian, S Hussain, J Guo, H Yu, and X Wang, "A brief survey of visual saliency detection," *Multimedia Tools Appl* **79**, 45–46, 34605–34645 2020
4. B Dedhia, J-C Chiang, and Y-F Char, "Saliency Prediction for Omnidirectional Images Considering Optimization on Sphere Domain," *IEEE Int Conference on Acoustics, Speech and Signal Processing - Proceedings* pp 2142–2146, 2019
5. F Zhang and T Wu, "Video salient region detection model based on wavelet transform and feature comparison," *J Image Video Proc* **2019**, 58, vol 2, 2019
6. M K Bashar, N Ohnishi, and K Agusa, "A New Texture Representation Approach Based on Local Feature Saliency," *Pattern Recognit Image Ana*, vol 17, no 1, pp 11–24, 2007
7. C He, Z Chen, and C Liu, "Salient object detection via images frequency domain analyzing," *Signal, Image Video Process*, vol 10, no 7, pp 1295–1302, 2016

8. İ Nevrez, W Lin, S Member, and Y Fang, "A Saliency Detection Model Using Low-Level Features Based on Wavelet Transform," *IEEE Trans Multimedia* vol 15, no 1, pp 96–105, 2013
9. X Hou and L Zhang, "Saliency Detection: A Spectral Residual Approach" *Proc IEEE Int Conf Comput Vision and Pattern Recognition*, 2007, pp 1–8
10. C Guo, Q Ma, and L Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform," *Proc IEEE Int Conf Comput Vision and Pattern Recognition*, 2008, pp 1–8
11. M Rezaei and A M Omair, "Salient region detection using efficient wavelet-based textural feature maps," *Multimed Tools Appl* 77 pp 16291–16317, 2018
12. L Zhang, J Chen, S Member, and B Qiu, "Region-of-Interest Coding Based on Saliency Detection and Directional Wavelet for Remote Sensing Images," *IEEE Geosci Remote Sens Lett* 2017, 14, 23–27 vol 14, no 1, pp 23–27, 2017
13. R Achanta, S Hemami, F Estrada, S Sabine, and D L Epfl, "Frequency-tuned Salient Region Detection," *CVPR*, pp 1597–1604 (2009) pp 1597–1604, 2009
14. A Aggoun and M Mazri, "Wavelet-based compression algorithm for still omnidirectional 3d integral images," *Signal, Image Video Process*, vol 2, no 2, pp 141–153, 2008
15. Y Zhang and Y Sun, "Optik An image watermarking method based on visual saliency and contourlet transform," *Opt - Int J Light Electron Opt*, vol 186, no February, pp 379–389, 2019
16. D Chen, T Jia, and C Wu, "Visual saliency detection: From space to frequency," *Signal Process Image Commun*, vol 44, pp 57–68, 2016
17. M Jian, W Zhang, H Yu, C Cui, X Nie, H Zhang and Y Yin, "Saliency detection based on directional patches extraction and principal local color contrast q," *J Vis Commun Image Represent*, vol 57, pp 1–11, 2018
18. A Teynor, H Burkhardt, and H B De, "Wavelet-based Salient Points with Scale Information for Classification," *Procs of the 19th Int Conf on Pattern Recognition*, Proc IEEE Int Conf 2008
19. P Zhang, W Liu, H Lu, and C Shen, "Salient Object Detection with Lossless Feature Reflection and Weighted Structural Loss," *IEEE Trans Image Process*, vol 28, no 6, pp 3048–3060, 2019
20. A De Abreu, C Ozcinar, and A Smolic, "Look around you: Saliency maps for omnidirectional images in VR applications," *2017 9th Int Conf Qual Multimed Exp QoMEX 2017*, pp 1–6, 2017
21. H Duan, G Zhai, X Min, Y Zhu, Y Fang, and X Yang, "Perceptual Quality Assessment of Omnidirectional Images," *Proc - IEEE Int Symp Circuits Syst*, vol 2018-May, pp 1–5, 2018
22. T Maugey, O Le Meur, and Z Liu, "Saliency-based navigation in omnidirectional image," *2017 IEEE 19th Int Work Multimed Signal Process MMSP 2017*, vol 2017-Janua, pp 1–6, 2017
23. F Y Chao, L Zhang, W Hamidouche, and O Deforges, "Salgan360: Visual saliency prediction on 360 degree images with generative adversarial networks," *2018 IEEE Int Conf Multimed Expo Work ICMEW 2018*, pp 1–4, 2018
24. R Dudek, S Croci, A Smolic, and S Knorr, "Signal Processing : Image Communication Robust global and local color matching in stereoscopic omnidirectional content," *Signal Process Image Commun*, vol 74, no March, pp 231–241, 2019
25. L Ye et al, "Saliency Detection Via Similar Image Retrieval," *IEEE Signal Process Lett*, vol 23, no 6, pp 838–842, 2016
26. M Longkumer and H Gupta, "Image Denoising Using Wavelet Transform, Median Filter and Soft Thresholding," *International Research Journal of Engineering and Technology*, 5(7) pp 729–732, 2018
27. X Ma, X Xie, K Man Lam, Y Zhong, "Efficient saliency analysis based on wavelet transform and entropy," *J Vis Commun Image Represent*, vol 30, pp 201–207, 2015

28. P Lebreton and A Raake, "Signal Processing : Image Communication GBVS360 , BMS360 , ProSal : Extending existing saliency prediction models from 2D to omnidirectional images," *Signal Process Image Commun*, vol 69, no September 2017, pp 69–78, 2018
29. Y Fang, X Zhang, and N Imamoglu, "A novel superpixel-based saliency detection model for 360-degree images," *Signal Process Image Commun*, vol 69, no August, pp 1–7, 2018
30. I Cinaroglu and Y Bastanlar, "A direct approach for object detection with catadioptric omnidirectional cameras," *Signal, Image Video Process*, vol 10, no 2, pp 413–420, 2016
31. V S Jabade, "Literature Review of Wavelet Based Digital Image Watermarking Techniques," *Int J Comput Appl*, vol 31, no 1, pp 28–35, 2011
32. Y Wang, X Zhao, X Hu, Y Li, and K Huang, "Focal Boundary Guided Salient Object Detection," *IEEE Trans Image Process*, vol 28, no 6, pp 2813–2824, 2019