Fast Visual Saliency Detection with Bisection Search to Scale Selection

Omprakash Rajankar
Mukesh Patel School of Technology
Management and Engineering,
NMIMS University
Mumbai, India, 400056
Email: online.omrajankar@gmail.com

Uttam Kolekar
Department of Electronics and
Telecommunication Engineering
P. Shah Institute of Technology
Thane, India
Email: uttamkolekar@gmail.com

***Abstract*—Inan image, objects are meaningful entity only at certain scales; however appropriate scales of object cannot be predicted a priori. The description of the image at multiple scales is thus necessary for salient objects detection. Multiple computations involved in this process make the model slow. To reduce the computational complexity bisection search to scale selection is proposed in this paper. The saliency detection with Hypercomplex Fourier Transform is made faster with the proposed algorithm. Minimum and maximum scales act as distinct seeds. To begin with, saliency maps and their entropy are obtained for these seeds. A set of lower half or upper half scales is selected recursively based on minimum entropy for further search. A saliency map of approximate minimum entropy is obtained efficiently. The proposed method reduces computational complexity from *O (N) to O* (log_2N+1) compared to the HFT method in the literature.**

***Keywords*—Bisection Search; Object scales; Quaternion Hypercomplex Fourier Transform; Scale Space Analysis; Visual Saliency**

I. INTRODUCTION

Objects in an image become visible differently based on the scale of observation[1]. For example what appears to be a green patch of land in satellite image turns out to be a forest in a closer look of the land from an airplane. An even closer look from the top of a building can show different parts; trees, branches and leaves. Texture of leaves can be observed under magnifying glass and so on. The green patch is observed at a coarse scale while the leaves are visible at the much finer scale. Analogous to objects in the real world, details in an image exist over a limited range of resolution[2]. Thus, the concept of scale is very much important while processing the image to highlight objects in it. The object scale and their spatial frequencies are also very much associated [3].

Visual salience is the state or distinct subjective perceptual quality which forces some items in the world that pop out from their adjacent items and quickly grabs our visual attention [4]. Such as a colorful rainbow in the sky that makes it pop out from its surroundings and draw our attention in an automatic and rapid manner. Similarly some parts of image are more interesting or visually important, that catch our visual attention. Thus, salient regions are described as uncommon or unique samples, of the image having distinct features. The regular samples of the image may then be termed as non-salient regions. Computing visual salience has been a topic of interest since last 25 years. Systematic scanning of images from left-to-right and top-to bottom is the traditional method of attempting to locate objects of interest[4][5],[6]. Whereas, visual salience provides relatively rapid mechanism to select a few most likely objects and overlook the clutter present in an image[4]. Top-down and bottom-up are the two approaches used to drive visual attention. Low level image features like orientation, color, intensity and contrast drives the bottom-up approach. Top down approach is task driven with prior knowledge and is much slower than bottom-up approach [7]. Many computational, spatial biological and hybrid models of the saliency detection are developed in last two decades. State of art visual saliency models are computationally complex and slow in saliency detection. Frequency domain transforms, specifically Fast Fourier Transform (FFT) are employed in computational methods of saliency detection. FFT of an image results in the amplitude and phase spectrum. All the frequency components present in the image are also present in its amplitude spectrum. Saliency map can be obtained by convolution of the amplitude spectrum with Gaussian kernel of a right scale, maintaining the other information unchanged[1].

Frequency domain saliency detection models are faster in detecting saliency than spatial domain models hence becoming popular. Still they are not fast enough to match the requirement of real time applications. This paper explores bisection search to scale space analysis to reduce the computational complexity significantly. To explain this concept an example of Hypercomplex Fourier Transform (HFT) based saliency model of [8]is considered in this paper.

The organization of this paper is as follows: section II is related to background, which discusses the relation between object scale and spatial frequency. It also describes Quaternion Hypercomplex Fourier Transform (QHFT). Section III explains the use of QHFT for visual saliency detection. Section IV presents the proposed bisection search for appropriate scale selection. For experimentation two databases 'ImgSal V1.0' and 'MIT Saliency Benchmark' are used. More than one evaluation scores are considered for performance evaluation. The result and comparison with state of art methods in the literature are shown in Section V. Section VI concludes with the utilization of the proposed approach.

978-1-4799-6272-3/15/$31.00(c)2015 IEEE

II. BACKGROUND

A. Object Scale and spatial frequency

The objects in an image become visible in different ways depending on its scale. There is close relation among the object scale and spatial frequency. Low frequencies represent large objects and background of the image; high frequencies define object boundaries, whereas highest frequencies specify noise, coding artifacts, and texture patterns. Medium frequencies generally represent the most important parts called as salient objects. Convolution of given image $I(x,y)$ with the Gaussian kernel $g(x,y;\sigma)$ $g(x,y;\sigma)$ is its Gaussian scale-space representation $L(x,y;\sigma)$, described in [9]as (1)

$$L(x,y;\sigma) = (g(x,y;\sigma) * I(x,y) \tag{1}$$

The 2-D Gaussian kernel is expressed as (2):

$$g(x,y;\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{2}$$

Where, σ is standard deviation of the Gaussian kernel. The object scale *k* and σ are expressed as (3) in [1].

$$\sigma = 2^{k-2} \tag{3}$$

Typically only a finite discrete set of levels of L for $\sigma \geq 0$ are considered in the scale-space representation. Fourier transform of Gaussian $F(u,v)$ is Gaussian, so it can generate linear scale-space representation.

$$F(u,v) = e^{-2\pi^2(u^2+v^2)\sigma^2} \tag{4}$$

As σ increases, L is the result of smoothing I with a larger Gaussian filter, thereby removing high frequency details. The object scale and cut-off frequency of low-pass or band-pass filters that are used in saliency detection are very much related. The higher cut-off frequency of the filters depends on how well defined boundaries are needed, and the amount of high frequency detail is to be discarded. The bandwidth of the filters should be narrow to highlight salient objects; at the same time it should suppress background, noise, coding artifacts, texture and repeated patterns. The use of band-pass filter can also control the lower cut-off frequency to narrow the bandwidth. The larger is the scale of the object; the lesser is the lower cutoff frequency f_{LC} of the band-pass filter in order to completely highlight it. For edge detection f_{LC} is kept high which narrows down the bandwidth [10]. Similarly, a narrow bandwidth is used for corner and interest point detection [11], [12].It is observed that detection of scale of large elongated object is difficult. To some extent this problem can be solved by segmenting the object if it is known a priori else all of the low frequency content should be used. In fact, there is no way to know a priori what scales are appropriate for describing the objects in a image, hence, descriptions of the image at multiple scales is considered for salient object detection in [8] and [13].This approach becomes computationally complex because of multiple scale analysis.

B. Quaternion Hypercomplex Fourier Transform

The traditional Discrete Fourier Transform (DFT) is very much useful tool to process gray scale images or color separated component of color images. There are some applications which need to utilize correlation between the color components. Hypercomplex Fourier Transform proposed in [14], [15]is found useful to process color images in total rather than with individual color components. Fourier transform operates on a real matrix. However, if more than one image features are combined into a hypercomplex matrix, then each element turns out to be a vector and the hypercomplex matrix results in a vector field. In such situation the traditional Fourier Transform proves unsuitable. For color images, to define the HFT, the hypercomplex numbers specifically quaternion are used. The quaternion formally introduced by Hamilton in 1843 is an associative noncommutative four-dimensional algebra. Quaternion is expressed as:

$$f(m,n) = q_1 + q_2 i + q_3 j + q_4 k \tag{5}$$

Where q_1 is scalar part and $q_2 i + q_3 j + q_4 k$ is a vector partin which $q_1, q_2, q_3, and\ q_4$ are real numbers and i, j, k are complex operators that obeys

$$i^2 = j^2 = k^2 = ijk = -1 \tag{6}$$

And

$$ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j, \tag{7}$$

The rules given above are consistent with the usual definition of the vector product, which is a part of the full quaternion product. Addition and multiplication are defined exactly analogously to complex addition and multiplication, but note that the operators may not be reordered arbitrarily as in the case of complex arithmetic. With this condition, implementation of quaternion arithmetic in hardware or software is straightforward. Quaternion with zero scalar part is termedas pure quaternion.

Given a hypercomplex matrix, the left sided discrete version of the HFT and inverse HFT are given by (8) and (9) respectively. Both $F_H[u,v]$ and $f(m,n)$ are hypercomplex matrix.

$$F_H[u,v] = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} e^{-\mu 2\pi \left(\left(\frac{nu}{N}\right) + \left(\frac{mv}{M}\right) \right)} f(m,n) \tag{8}$$

$$f(m,n)=\frac{1}{\sqrt{MN}}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}e^{+\mu 2\pi\left(\left(\frac{nu}{N}\right)+\left(\frac{mv}{M}\right)\right)}F_H[u,v] \quad (9)$$

Where, μ is unit pure quaternion and $\mu^2=-1$ [8].

III. VISUAL SALIENCY USING QHFT

A. Representation of Feature Maps

The features considered for saliency detection in the HFT model of [8] are intensity and two opponent color space representations. The saliency model can be extended to incorporate motion as a fourth feature vector of quaternion for saliency detection in video. Feature maps f_2 , f_3 and f_4 are computed from the three color channels; r,g,b of the image as.

$$f_2=\frac{r+g+b}{3} \quad (10)$$

$$f_3=\left(r-\frac{g+b}{2}\right)-\left(g-\frac{r+b}{2}\right) \quad (11)$$

$$f_4=\left(b-\frac{r+g}{2}\right)-\left(\frac{r+g}{2}-\frac{|r-g|}{2}-b\right) \quad (12)$$

The Hypercomplex matrix *f (m, n)* is described by combining these feature maps according to (13);

$$f(m,n)=w_1f_1+w_2f_2i+w_3f_3j+w_4f_4k \quad (13)$$

Where $w_1,w_2\,w_3,w_4$ are the weights and f_1,f_2,f_3,f_4 are the feature maps;f_1 may be used for motion feature,$f_1 = M$, and f_2,f_3,f_4 as three features of the image. For the static image weight of motion feature is set as $w_1=0$. Weights of image features like intensity f_2 and opponent colors f_3 and f_4 are set as $w_2=0.5,w_3=0.25,w_4=0.25$ respectively[8].

B. Computation of Saliency Map

Polar form representation of $F_H[u,v]$ in (8) is as follows:

$$F_H[u,v]=\|F_H[u,v]\|e^{(\mu\phi(u,v))} \quad (14)$$

Where $\|F_H(u,v)\|$, $\phi(u,v)$ and $\mu(u,v)$ are called as amplitude spectrum $A(u,v)$, phase spectrum $\phi(u,v)$ and a pure quaternion matrix the eigenaxis spectrum respectively.

The amplitude spectrum contains both salient and non-salient information about the scene. The smooth spectrum scale space $\Delta = \{\Delta_k\}$is created by convolving $A(u,v)$with Gaussian kernels according to (15) [8].

$$(u,v;k)=\left(g(.,.;k)*A\right)(u,v) \quad (15)$$

Saliency maps S_k can now be obtained at each scale *k*by performing inverse transform on smoothed amplitude spectrum Δ_k, and the original phase and eigenaxis spectra, by (16):

$$S_k=g*F_H^{-1}\left\{\Lambda_k(u,v)e^{\mu\phi(u,v)}\right\}^2 \quad (16)$$

Where, Gaussian kernel g is of specific scale. In this way a series of saliency maps S_k is obtained. Fig. 1, shows Gaussian filtering of a 128x128 Image at scales *k = 1 to 8*. It reveals that when the scale k is very small, the information contained in the amplitude plots is retained quite well, but when it becomes very large, the pertinent information is lost[8]. In a saliency map, salient regions are in highlighted form and rest of the map is in suppressed form. So the histogram of saliency map is expected to be clustered around certain values and its entropy would be very small. With this assumption, in the HFT saliency model of [8] the final saliency map *S* is chosen from the set of saliency maps $\{S_k\}$ for specific scale based on minimum entropy.

IV. BISECTION SEARCH TO SCALE SPACE

As mentioned earlier to describe object predicting its scale a priori is difficult. Hence, description of the image at multiple scales is necessary. In HFT saliency detection model convolution of amplitude spectrum of a given image with a series of Gaussian kernels produce a set of saliency maps. One map is obtained per scale. The size of image decides the maximum number of scales.

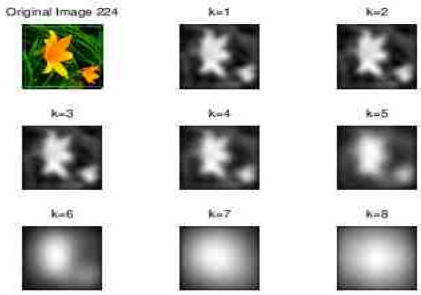

Fig. 1. Gaussian filtering of a image at scales 1 to 8

For example, for 128x128 image size, maximum value of scale *N* is eight. For fixed number of scales given image is required to be resized. Generally, eight or sixteen scales are used for the scale space analysis. Multiple scales analysis in HFT saliency detection makes the model computationally

complex. Local minimum bisection search is proposed to solve this problem. Bisection search starts with minimum 1, and maximum N, values of scales as seed points. Saliency maps and their entropy are obtained for these scales. A set of lower half or upper half scales is selected recursively based on minimum entropy for further search. Finally, a saliency map with local minimum entropy is obtained. Instead of finding saliency maps for all the scales, the algorithm finds only those saliency maps which are required for comparison, in minimum entropy search path of the algorithm. Sometimes the local minima trap the bisection search. To overcome this situation and to search beyond the neighboring small rise or hill in entropy, care may be taken. The proposed method reduces computational complexity from *O (N)* to*O (log₂N + 1)* as compared to the methods in the literature. This coding algorithm incorporate bisection search in the 4th step of HFT coding algorithm of Jian Li *et.al.*[8].

A. *Codingalgorithm*

Algorithm: QHFT Visual saliency model with Bisection search

Input: The resized color image *C* (*m, n)*

Output: Saliency map *S (m, n)* of*C (m,n)*

1) Calculate the feature maps $\{f_2, f_3, f_4\}$ of *C(m,n)*according to (10-12);
2) Combine these features to obtain the Hypercomplex matrix *f(m, n)* according to (13);
3) Compute amplitude spectrum of $f(m,n)$ by taking its HFT. Preserve the exponential term which consists of phase and eigenaxis spectrum.
4) Use Gaussian kernels to smooth the amplitude spectrum according to (15) and obtain saliency maps according to (16)for specific scale selected by the following heuristic approach. Let $k = 1, 2, 3..N$ where *N* is the maximum scale obtained from image size. Set number of iterations equal to $log_2 N + 1$, and seed scale values of k_l equals to 1 and k_h equals to *N*.
   i. Derive Saliency maps and calculate entropy for scales 1 and *N*. Repeat following steps (ii) and (iii) for the set value of iterations.
   ii. If entropy for scale 1 is smaller than entropy for scale *N*, replace scale k_h by $(k_l + k_h)/2$ else replace scale k_l by $(k_l + k_h)/2$.
   iii. Obtain Saliency map and calculate saliency for the new scale.
   iv. Final saliency map *S* is the saliency map corresponding to minimum entropy.
5) Return *S*.

V. Experimentation and Results

To evaluate the performance of proposed saliency model, we compare saliency maps generated by it and other state of art saliency models such as Itti[16], HFT[8] and AC [13] with the human-labeled salient regions called as ground truth. Test images and respective ground truth are from the ImgSal V1.0 database by Jian Li[17] which has about 235 images of different category. Third party evaluation is also done with the MIT Saliency Benchmark database [18]which has about 300 benchmark images. A perfect ground truth fixation map would come from an infinity observers. MIT Saliency benchmark is created from 39 viewers has a performance of 0.899 which captures about 95% of ground truth, and are assumed as accurate [19]. The ground truth of this database are not public, hence the saliency maps of proposed model are evaluated by them only.

For analyzing saliency models, more than one evaluation scores are considered to make the conclusions independent of the choice of metric. The evaluation of a saliency model is done by comparing the estimated saliency map *S* and the ground truth map *G*.

The Area Under the ROC Curve (AUC) is calculated according to [20]. AUC indicates how well the saliency map predicts actual human eye fixations. AUC score of one indicate correct prediction whereas a score of 0.5 corresponds to chance level [21]. Fig. 2 shows the average Receivers Operating Characteristic (ROC) curves for AC, Itti, HFT and proposed HFT-bisection model. Database used for evaluation is 'ImgSal V1.0 dataset'. It is observed that the area under the curve is largest, for HFT and the proposed model.

Linear Correlation Coefficient (CC) measures the strength of linear correlation between two variables. A CC score of 1 indicates perfect linear relationship whereas a score of 0 means the two variables are not correlated [22].

The Spectrum Scale Space (SSS) is the average value of the Normalized Scan-path Saliency map (NSS) at human eye fixation locations [23]. A NSS score of 1 indicates that the subject look into the region whose predicted saliency is more than average by one standard deviation, whereas a score of 0corresponds to a chance in predicting humane

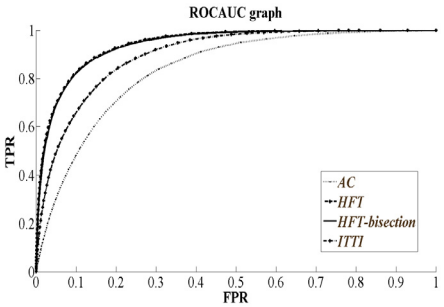

Fig. 2. Receivers Operating Characteristic (ROC) curves

S larity between two different saliency maps when viewed as distributions [24]. It is the sum of the minimum values at each point in the distributions. A SIM score of 1 indicate

that the distributions are identical whereas a score of 0.5 correspond completely different distributions.

The Earth Mover's Distance (EMD) is a measure of the distance between two probability distributions over a region. [23] used faster implementation of EMD without a threshold provided by [25]. EMD computes the minimal cost to transform the probability distribution of the saliency maps into the one of the human eye fixations. An EMD score of 0 indicate that the distributions are identical whereas a score of 1 correspond to no overlap and distributions are completely different. The Peak value of the DSC curve (PoDSC) corresponds to the optimal threshold and the best possible algorithm performance [26].

Table I list the average evaluation scores AUC, PoDSC, CC, NSS, SIM and time to obtain saliency map of resized color test images of128x128 size on 3.20GHz i5-CPU. The evaluation scores of proposed method are close to that of HFT and out performs the others. The time required for saliency detection is reduced as compared to HFT.

TABLE I. EVALUATION SCORES FOR DATABASE IMGSAL

Model	AUC	PoDSC	CC	NSS	SIM	Time to compute Saliency map (s)
Achanta(AC)	0.83	0.41	0.46	0.13	0.60	0.032
HFT	0.94	0.59	0.67	0.16	0.635	0.234
HFTbisection	**0.94**	**0.58**	**0.66**	**0.16**	**0.63**	**0.126**
Itti	0.89	0.46	0.52	0.13	0.58	1.1

The proposed model is also evaluated by third party on their own database: MIT Saliency Benchmark database [18]. The results are listed for comparison in Table II for the four models under consideration. The performance scores of proposed model are found better than state of art models. It is observed that the performance scores of proposed model and HFT are matched for this database also.

TABLE II. EVALUATION WITH MIT SALIENCY BENCHMARK

Model	AUC-1	EMD	CC	NSS	SIM
baseline: infinite humans	0.91	0	1	3.18	1
HFT	0.76	3.78	0.39	1.06	0.45
HFTbisection	**0.76**	**3.78**	**0.39**	**1.06**	**0.45**
Itti-Koch	0.60	5.17	0.14	0.43	0.20
Achanta(AC)	0.52	5.77	0.04	0.13	0.29
Baseline: Chance	0.50	5.73	0.00	0.00	0.31

VI. CONCLUSION

Scale of salient object is an important factor but it is not known a priori. This paper proposed bisection search to select appropriate scale for visual saliency detection. The algorithm only generates those saliency maps which fall on the search path. Therefore the proposed method reduced computational complexity from $O(N)$ to $O(\log_2 N + 1)$ as compared to the HFT method. For example for 16 scales, HFT method generates 16 saliency maps and selects one out of them, whereas the proposed method needs to generate only 5 maps. So the proposed model is faster than HFT model. As performance scores of proposed model and state of art HFT model are matched with each other it concludes that the salient object scale selection by the proposed method is as accurate as HFT model. Fast saliency detection makes the model appropriate for real-time applications, spatially for ROI based image compression.

ACKNOWLEDGMENT

The authors wouldlike to thank Z. Bylinskii, T. Judd, F. Durand, A. Oliva, and A. Torralba for running proposed saliency model on MIT Saliency Benchmark and providing the results for comparison with other state of art saliency models.

REFERENCES

[1] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at di erent scales," *J. Appl. Stat.*, 1994.

[2] P. Groot, C. Gilissen, and M. Egmont-Petersen, "Error probabilities for local extrema in gene expression data," *Pattern Recognit. Lett.*, vol. 28, no. 15, pp. 2133–2142, Nov. 2007.

[3] A. Witkin, "Scale-space filtering: A new approach to multi-scale description," in *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1984, vol. 9, pp. 150–153.

[4] Laurent Itti, "Visual salience - Scholarpedia," *Scholarpedia*, vol. 2, no. 9, p. 3327, 2007.

[5] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention.," *Vision Res.*, vol. 40, no. 10–12, pp. 1489–506, Jan. 2000.

[6] V. Navalpakkam and L. Itti, "Modeling the influence of task on attention.," *Vision Res.*, vol. 45, no. 2, pp. 205–31, Jan. 2005.

[7] A. Oliva, A. Torralba, M. S. Castelhano, and J. M. Henderson, "Top-down control of visual attention in object detection," in *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, 2003, vol. 1, pp. I–253–6.

[8] J. Li, M. D. Levine, X. An, X. Xu, and H. He, "Visual Saliency Based on Scale-Space Analysis in the Frequency Domain.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 1, Jul. 2012.

[9] R. Achanta, "Finding Objects of Interest in Images using Saliency and Superpixels," École Polytechnic Federale de Lausanne, 2011.

[10] P. P. Jonathan Harel, Christof Koch, J. Harel, C. Koch, and P. Perona, "Graph-Based Visual Saliency," *Nat. Rev. Immunol.*, Jun. 2006.

[11] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.

[12] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–118, 2004.

[13] R. Achanta, S. Hemami, F. Estrada, S. Sabine, and D. L. Epfl, "Frequency-tuned Salient Region Detection," *CVPR*, no. Ic, 2009.

[14] T. A. Ell, "Quaternion-Fourier transforms for analysis of two-dimensional linear time-invariant partial differential systems," in *Proceedings of 32nd IEEE Conference on Decision and Control*, 1993, pp. 1830–1841.

[15] T. A. Ell, S. Member, and S. J. Sangwine, "Hypercomplex Fourier Transforms of Color Images," vol. 16, no. 1, pp. 22–35, 2007.

[16] 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, 1998.

[17] J. Li, M. Levine, X. An, and H. He, “ImgSal: A benchmark for saliency detection v1.0,” in *Proceedings of the British Machine Vision Conference*, 2011, pp. 86.1–86.11.

[18] B. Zoya, T. Judd, D. Do, A. Oliva, and A. Torralba, “MIT Saliency Benchmark.” .

[19] T. Judd, F. Durand, and A. Torralba, “A Benchmark of Computational Models of Saliency to Predict Human Fixations A Benchmark of Computational Models of Saliency to Predict Human Fixations,” 2012.

[20] N. D. B. B. Bruce, J. K. Tsotsos, N. Bruce and J. Tsotsos, and J. K. Tsotosos, “Saliency Based on Information Maximization,” *J. Vis.*, vol. 7, no. 9, pp. 155–162, Mar. 2006.

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

[22] T. Jost, N. Ouerhani, R. Von Wartburg, R. Müri, and H. Hügli, “Assessing the contribution of color in visual attention,” *Comput. Vis. Image Underst.*, vol. 100, no. 1–2, pp. 107–123, Oct. 2005.

[23] R. J. Peters, A. Iyer, L. Itti, and C. Koch, “Components of bottom-up gaze allocation in natural images.,” *Vision Res.*, vol. 45, no. 18, pp. 2397–416, Aug. 2005.

[24] Z. Liu, O. Le Meur, and S. Luo, “Superpixel-based saliency detection To cite this version:,” 2013.

[25] O. Pele and M. Werman, “Fast and robust Earth Mover’s Distances,” in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 460–467.

[26] T. Veit, J. Tarel, P. Nicolle, and P. Charbonnier, “Evaluation of Road Marking Feature Extraction,” in *2008 11th International IEEE Conference on Intelligent Transportation Systems*, 2008, pp. 174–181.