An Efficient Region Detection base on Guided Filtering and Features Extraction

Sonali Karan
Dept. of Computer Science & Engineering
Samrat Ashok Technological Institute
Vidisha India

Shailendra Ku. Shrivastava, PhD
Dept. of Computer Science & Engineering
Samrat Ashok Technological Institute
Vidisha, India

ABSTRACT

Here in this paper a new and efficient technique for the Saliency Region Detection using Guided Filtering & Random Walker Segmentation for the Extraction of features from Natural Images and finally Support Vector Machines is applied for the training of these features. The Proposed Methodology implemented here is applied on MSRA Natural Image Datasets and comparison is done with the existing SGNR algorithm. The Proposed methodology provides higher Precision and recall rate in comparison.

Keywords

Saliency Region Detection, Natural Images, Guided Filtering, Random Walker Segmentation, SVM.

1. INTRODUCTION

Visual saliency is the perceptual qualities which make some items in the scene pop out from their surroundings and immediately attract our attention. It is well-known that humans can detect salient regions effortlessly even in highly clutter and complex scenes. The process of filtering out irrelevant and redundant visual information and detecting the most relevant parts of an image can significantly reduce the complexity of visual processing. An effective computational model for automatically generating saliency maps from images is thus of great interest to the computer vision community because it can facilitate many important computer vision and graphics applications. Many of these methods claim performance improvement over others; however, due to the difference in the principles used and the implementation details, the underlying relations among these methods remain hard to understand and their fundamental capabilities are unclear. To better understand the progress of this field, some recent survey papers have been published to qualitatively discuss the state-of-the-art methods on bottom-up saliency detection [1] and visual attention modeling [2, 3].
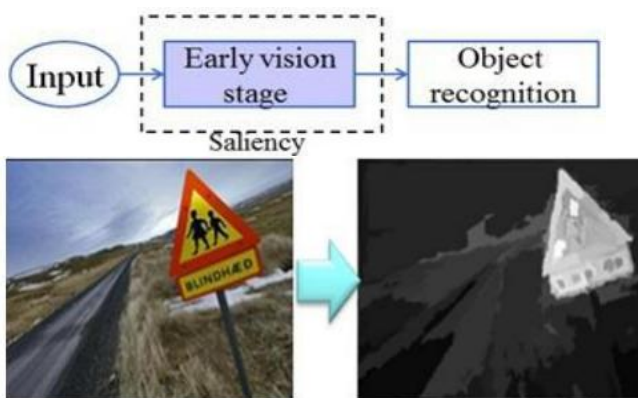

Figure-1: Illustration of saliency detection

Humans can spot familiar people or objects in a big crowd almost instantaneously, because (hypothetically) a colony of neurons would be responsible of detecting and recognizing them. Those familiar objects or shapes trigger what is called the "attention" mechanism, attention being obviously visual attention. In terms of visual attention, the familiar object is a salient feature in comparison with the rest of the image. Recent advances in hardware architecture and processing power gave rise to industrial interests in computer vision applications and algorithms. More specifically, there has been a significant progress in the field of object detection, which typically consists of a number of computationally intensive stages. It is believed that the human biological visual system has the early stage of attention [4], in which the background or clutter is discarded without the need of an exhaustive analysis of the scene as illustrated in Figure-1.This motivates many researchers to model the visual attention and propose saliency detection approaches that can provide an input to the object detection approaches. An observation that can be made with regards to saliency in images is that it is closely related to visual uniqueness in image attributes such as colors, edges, and boundaries. For example, in scenes such as that shown in Figure-1, the salient objects of interest (e.g., the traffic sign) exhibit very different visual characteristics than the rest of the scene i.e. the road and the background. From an information-driven perspective, one can also say that the object of interest is unique as it has low information non-redundancy within the given image compared to the rest of the scene, which is highly redundant. At the same time, it has high information content by being more complex in texture and edge shape in comparison to the rest of the scene.

Over the years, many methods have been devised to predict visually salient information in images which may also correspond to highly informative regions in the image. Existing saliency approaches can be sub-divided into purely bottom-up, top-down, and hybrid approaches combining top-down and bottom-up modes [5]. While bottom-up approaches compute saliency using image-based conspicuities, top-down approaches rely on a-priori known information about the appearance of target objects, tasks, and high-level cues to guide saliency towards important image regions.

Traditionally these salient regions have always been considered as local phenomena in which the salient regions stand out as local extrema with respect to their immediate neighbors. However, some recent advances in computer vision literature have indicated that these salient features can also be considered with respect to the entire image (or data ensemble) and can also be obtained via discriminate processes. However, due to the high-dimensional nature of image features, such an inference is not always feasible. Saliency can be defined as the state or quality of an object to

stand out relative to its neighbors [5]. A computer algorithm can be trained to determine the most visually salient regions in a top-down manner by recognizing the various objects present in the image first and then inferring if they can trigger visual attention. Alternatively, saliency can be inferred in a bottom-up manner completely diagnostic to the objects present in the image. In summary the features discussed in this chapter vary in the detail level they extract from images (See figure 2.1). The saliency detection algorithm presented in this thesis extracts distinct patterns which are least like other patterns in the image. Higher level features are more holistic than the low level features which represent local saliencies.
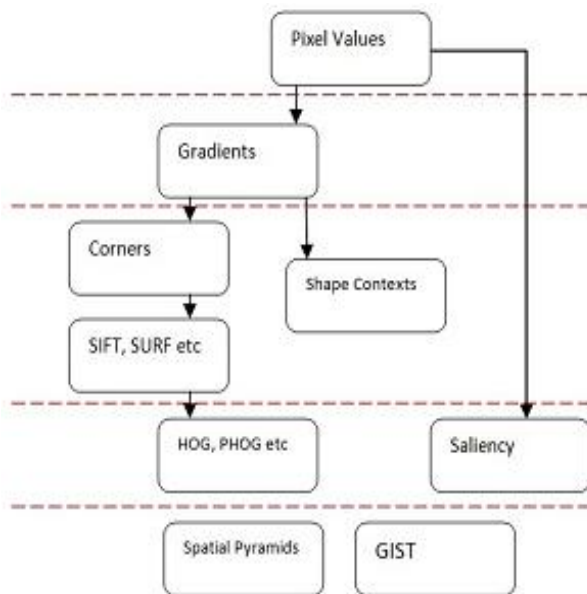

Figure 2: Hierarchy of image features.

A straightforward visual saliency based segmentation algorithm can be calculated be appropriating thresholding on the saliency map or probabilistic map. Several visual saliency based algorithms have been proposed in literature [6], [7]. Since the screening mammograms are gray scale images, algorithms which utilize the multi-channel properties of color images may not be useful.

2. LITERATURE SURVEY

Christian Scharffenberger et al. have proposed [8] an efficient structure for salient region detection in natural images based on the idea of self-guided statistical non-redundancy (SGNR). Salient regions are exceptional for the reason that they have low information redundancy within a given image at the same time as the break of the sight may extremely be unnecessary. We initial examine the structural features of the image using structured image elements (examples) and categorize them as being non-redundant or redundant based on textural compression and in general non-redundancy. This shows saliency detection toward regions with low information redundancy by allowing for unambiguously high information redundancy of samples potentially is in the right placing to the environment. Here then calculate the saliency map by formative the statistical non-redundancy of each example using a restricted graph model. Experimental outcomes supported on widely accessible data sets demonstrate that SGNR make available capable results when evaluated with accessible saliency methods.

In this paper [9], author has proposed a new saliency detection model is initiated by developing low-level features

acquired from the wavelet transform domain. Firstly, wavelet transform is utilized to generate the multi-scale feature maps which can characterize unusual features from edge to texture. The proposed model plans to transform local contrast at a place with its global saliency calculated found on the possibility of the features, and the proposed model think about local center-enclose differences and global contrast in the final saliency map.

Thus, the proposed model [9] can be used for images with different object sizes and levels of consistency. It also produces the saliency map with the same resolution of the input image. Another improvement is that the wavelet decomposition in the proposed model maintains awaiting the achievable coarsest amount is accomplished; with the intention of accomplish the feature maps. By this technique, they can acquire salient regions independent from their sizes or consistency due to the features with high contrast from edge to texture since the salient points can be described as the feature variations of the place with its environment supported on the WT coefficients. Experimental estimation shows the capable consequences from the proposed model by do better than the appropriate condition of the ability saliency detection models.

Liu et al. [10] removed multi-scale contrast, center-contain histogram and color spatial distribution feature maps from the image. These features can be joined in many techniques to acquire a saliency map. One probable method is to give identical weightage to all the three features. On the other hand, there can be an image which is salient in expressions of only singleton feature or grouping of two features with different influences or arrangement of all the three features with different weights. So a suitable weight vector is necessitated with the intention of join the features best possible for producing salient objects.

In this paper, this best possible weight vector for a given image is computed using C-PSO. The C-PSO make the most of a fitness function, which in our case is given in expressions of two parts; one of the modules is characterized in expressions of the consideration pixels and the additional in expressions of the background pixels. For object to be salient, the fitness element achieved in expressions of saliency assessments of attention pixels should be exploited and the other fitness part in expressions of saliency values of background pixels be reduced. The optimal weight vector so determined is used to combine the features into the final saliency map.

In this paper [11], author has propose a new quaternion illustration of an image and increased a multiresolution spatiotemporal saliency detection model called phase spectrum of quaternion Fourier transform (PQFT) to calculate the spatiotemporal saliency map from the image's quaternion demonstration. Each pixel of the image is characterized by a quaternion that consists of color, intensity and motion feature. Using the phase spectrum of QFT is used to determine the spatiotemporal saliency map, which thinks about not only salient spatial features like color, orientation and others in a single structure but also temporal feature among frames like motion. Their PQFT model is self-determining of earlier information and constraints, and experimental outcomes show that it is quick an adequate amount of to assemble real-time conditions (less than 1 ms per frame in C/C++ implementation), and do better than other modern models in computer vision. Since the suggested PQFT model can development the image under different resolutions, they commence the hierarchical selectivity (HS) structure based on

the PQFT model, which can be used to build a tree structure illustration of an image. A new multi-resolution wavelet domain foveation (MWDF) model is offered based on this tree structure, which can get better coding efficiency in image and video compression.

In [12], Achanta et al. deliberated that the saliency map should have definite borders, uniformly emphasizing the object if it is salient and most of all; the saliency map should be in high resolution. In the authors' estimation, with no be conventional to the points declared [12], the saliency would have maximum value effectiveness in definite applications. Consequently, Achanta et al. proposed a technique which produces the saliency map only by contrast representation. In their technique, the original image is primary smoothed using a 3-by-3 Gaussian filter to eradicate the high frequency content in the image. Then, the mean of the smoothed image is calculated. Finally, the saliency map is get hold of by the subtracting each pixel value of the original image from the calculated mean. while the approach used by Achanta et al. gives high resolution maps which has its effectiveness in several applications but in many other applications such as content based image retrieval (CBIR) all that substances is the detection of salient objection with suitable resolution. Actually, the come within reach of used in [12] will remove many small featured objects and textures which could be of consequence when the smoothing is functional to the spatial domain. Additionally, on condition that the saliency map make available the correct location of significant objects and is of sensible resolution (object can be visually recognized) the map can be think about adequate.

In another advance Discrete wavelet transform was used by Christopher et al. [13] to calculate the saliency map. This technique is doing well in detecting salient regions in an image with satisfactory resolution but it think about only the contrast of LL band; it completely take no notice of the other three detail bands. This perform will remove the smaller & finer features. But at times some significant information may be present in these feature constituents. In this paper they have suggested a process to calculate the saliency map think about the estimate in addition to featured coefficients of the discrete wavelet transform so that all the essential items can be emphasized & detected.

3. PROPOSED METHODOLOGY

The planned tactic implemented here consists of Following Stages for the Recognition and Classification of Natural Images & Region Detection.

1. Take an input Training Dataset of Natural Images.

2. Apply Guided Filtering the Input Dataset.

3. Apply Random Walker Segmentation on the Image.

4. Extract features from the image.

5. Apply Support Vector Machine to Train the Extracted Features.

6. Region Detection based on Classified Features.

Guided Filtering
It is based on a general linear translation-variant filtering process, which involves a guidance image I, an input image p, and an output image q. Both I and p are given beforehand according to the application, and they can be identical. The filtering output at a pixel i is expressed as a weighted average:

$$q_i = \sum_j W_{ij}(I)p_j$$

Where 'i' and 'j' are the pixel indexes. The Filter Kernel Wij is a function of the guidance image I and independent of p. This filter is linear with respect to p.

The guided filtering can be applied using the following steps:
1. Take an input gray scale image.

2. It takes input parameters as r and esp., since it works on the local window and regularization parameters.

3. Take the mean of the image, parameter P and I*P.

4. Then covariance of the Image I and P can be computed.

5. Take the mean of I*I.

6. Apply filter on the parameters a and b and take the mean of the image for a and b.

Support Vector Machine
Consider training sample$\{(x_i, d_i)\}$, where x_i is the input pattern, d_i is the desired output:

$$W_0^T X_i + b_0 \geq +1, for\ d_i = +1$$
$$W_0^T X_i + b_0 \leq -1, for\ d_i = -1$$

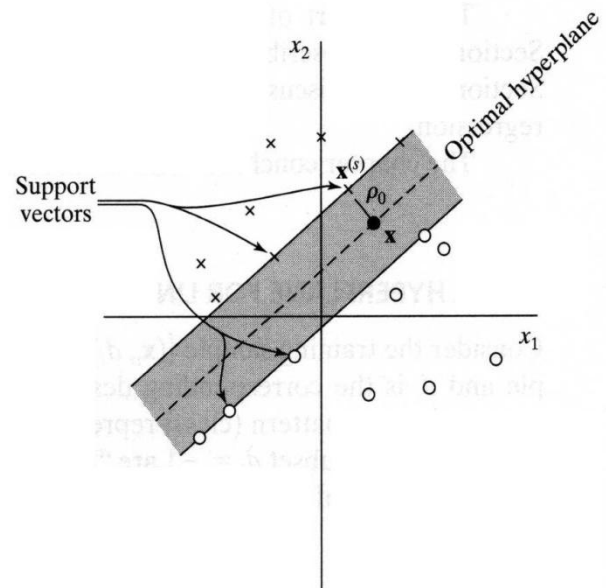**Figure 3. Basic Architecture of SVM**

The data point which is very near is called the margin of separation ρ

The main aim of using the SVM is to find the particular hyperplane of which the margin ρ is maximized

Optimal hyper plane $W_0^T X + b_0 = 0$

For example, if we are choosing our model from the set of hyperp lanes in Rn, then we have:

$$f(x;\ \{w;\ b\}) = sign(w\ .\ x + b)$$

We can try to learn *f(x; _)* by choosing a function that performs well on training data:

$$R_{emp}(\alpha) = \frac{1}{m}\sum_{i=1}^{m} l(f(x_i, \alpha), y_i)$$

Random Walker Segmentation

It is a technique of segmentation on the basis of selecting foreground and background as seed pixels by moving randomly to other pixels moving from background till any foreground pixel is obtained and the region is extracted as segmented region from the image.

4. RESULT ANALYSIS

The Table Shown below is the analysis and comparison of Precision rate on Natural Image Dataset. The Comparison is done between the existing work and the planned methodology implemented. The Experimental Result Analysis and Comparison is done for a number of Samples from the Natural Image Dataset for 50, 100, 150, 200, 250 and 300 Natural Images and the planned procedure implemented provides better and improved Precision Rate as compared to the existing Methodology implemented for Region Detection.

Table 1. Analysis of Precision

	Precision	
# of Samples	Existing Work	Proposed Work
50	0.925	0.94
100	0.93	0.948
150	0.934	0.952
200	0.94	0.956
250	0.95	0.972
300	0.956	0.977

The Table Shown below is the analysis and comparison of Recall rate on Natural Image Dataset. The Comparison is done between the existing work and the planned procedure implemented. The Experimental Result Analysis and Comparison is done for a number of Samples from the Natural Image Dataset for 50, 100, 150, 200, 250 and 300 Natural Images and the planned procedure implemented provides better and improved Recall Rate as compared to the existing Methodology implemented for Region Detection.

Table 2. Analysis of Recall

	Recall	
# of Samples	Existing Work	Proposed Work
50	0.6	0.68
100	0.65	0.73
150	0.7	0.79
200	0.75	0.86
250	0.8	0.89
300	0.85	0.93

The Table Shown below is the analysis and comparison of Execution Time on Natural Image Dataset. The Comparison is done between the existing work and the planned procedure implemented. The Experimental Result Analysis and Comparison is done for a number of Images from the Natural Image Dataset for different patch sized and the planned procedure implemented provides better and improved Execution Time as compared to the existing Methodology implemented for Region Detection.

Table 3. Analysis of Execution Time

	Execution Time (S)	
Patch Size	Existing Work	Proposed Work
3*3	1.097	0.95
5*5	1.773	1.34
7*7	3.21	2.87
9*9	5.614	5.25
11*11	8.735	8.42
13*13	14.121	13.93

The Figure Shown below is the analysis and comparison of Precision rate on Natural Image Dataset. The Comparison is done between the existing work and the planned procedure implemented. The Experimental Result Analysis and Comparison is done for a number of Samples from the Natural Image Dataset for 50, 100, 150, 200, 250 and 300 Natural Images and the planned procedure implemented provides better and improved Precision Rate as compared to the existing Methodology implemented for Region Detection.
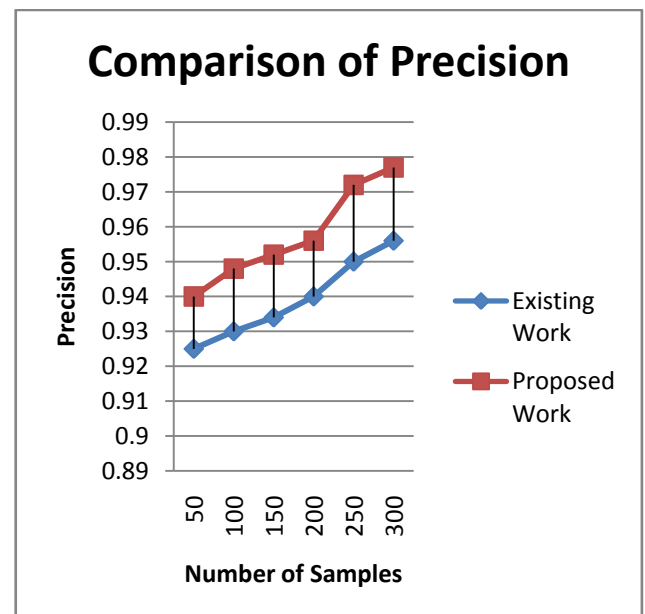
Figure 4. Comparison of Precision

The Figure Shown below is the analysis and comparison of Recall rate on Natural Image Dataset. The Comparison is done between the existing work and the planned procedure implemented. The Experimental Result Analysis and Comparison is done for a number of Samples from the Natural Image Dataset for 50, 100, 150, 200, 250 and 300 Natural Images and the planned procedure implemented provides better and improved Recall Rate as compared to the existing Methodology implemented for Region Detection.
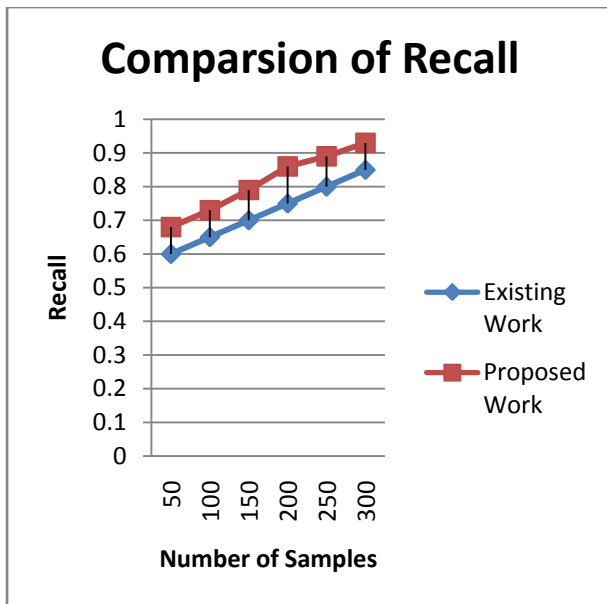
Comparsion of Recall

Figure 5. Comparison of Recall

The Figure Shown below is the analysis and comparison of Execution Time on Natural Image Dataset. The Comparison is done between the existing work and the planned procedure implemented. The Experimental Result Analysis and Comparison is done for a number of Images from the Natural Image Dataset for different patch sized and the planned procedure implemented provides better and improved Execution Time as compared to the existing Methodology implemented for Region Detection.
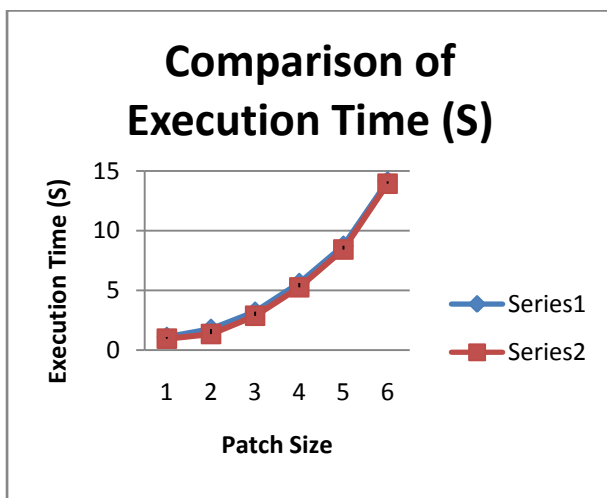
Comparison of Execution Time (S)

Figure 6. Comparison of Execution Time

5. CONCLUSION AND FUTURE WORK

The Proposed Methodology implemented for the Region Detection in Natural Images using Filtering of Natural Images using Guided Filtering and Segmentation using Random Walker Segmentation & Features Extraction using Support Vector Machine provides efficient results in Comparison with the existing methodology implemented for Natural Images.

The Experimental results are performed on MSRA Dataset of Natural Images in which the proposed methodology provides efficient Precision & Recall as compared to the SGNR Algorithm.

Although the technique implemented here is feasible and efficient as compared to the existing Recognition and Classification of Natural Images but there are some future directions that container be implemented further.

1. Improvement of Accuracy rate.

2. Classification and Recognition of Natural Images on other datasets.

3. Improvement of Computational time.

6. REFERENCES

[1] A. Toet, "Computational versus psychophysical image saliency: A comparative evaluation study," PAMI, vol. 99, no. 1, 2011.

[2] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," PAMI, 2012.

[3] S. Frintrop, E. Rome, and H. Christensen, "Computational visual attention systems and their cognitive foundations: A survey," ACM Transactions on Applied Perception (TAP), vol. 7, no. 1, p. 6, 2010.

[4] M. Carandini, J. Demb, V. Mante, D. Tolhurst, Y. Dan, B. Olshausen. "Do we know what the early visual system does?" Journal of Neuroscience, vol. 25, pp.10577-10597, 2005.

[5] L. Itti and C. Koch "Computational modeling of visual attention" Nature Reviews Neuroscience, 2(3):194-203, 2001.

[6] BORJI, A., AND ITTI, L. State-of-the-Art in Visual Attention Modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence 35, 1 (2013), 185–207.

[7] BORJI, A., SIHITE, D. N., AND ITTI, L. Quantitative Analysis of Human-Model Agreement in Visual Saliency Modeling: A Comparative Study. IEEE Transactions on Image Processing 22, 1 (2013), 55–69.

[8] C. Scharfenberger et al.: Salient Region Detection Using SGNR IEEE, 2015 DO1-0.1109/ACCESS.2015.2502842.

[9] Nevrez İmamoğlu, Weisi Lin, and Yuming Fang, "A Saliency Detection Model Using Low-Level Features Based on Wavelet Transform" IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 1, JANUARY 2013

[10] T. Liu, Z. Yuan, J. Sun-Wang, N. Zheng, X. Tang, H.Y. Shum, Learning to detect a salient object, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 353–366.

[11] Chenlei Guo, and Liming Zhang, "A Novel Multiresolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 1, JANUARY 2010,

[12] R. Achanta, S. Hemami, F. Estrada, and S. Sllsstrunk, "Frequency tuned Salient Region Detection", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[13] C. Ngau, Li-Minn and K. Seng "Bottom up Visual Saliency Ma using Wavelet Transform Domain" 978-1-4244-5540-9/10/$26.00 IEEE, 2010.